(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

CARL MURCHISON

Gordon Allport
Harvard University

Arthur G. Bills
University of Cincinnati

E. A. Bott
University of Toronto

Leonard Carmichael
Smithsonian Institution

Herbert S. Conrad
U. S. Office of Education

Elmer K. Culler
University of Rochester

John F. Dashiell
University of North Carolina

Roy M. Dorcus
University of California at Los Angeles

Horace B. English
Ohio State University

Franklin Fearing
University of California at Los Angeles

Frank A. Geldard
University of Virginia

Clarence H. Graham
Columbia University

Ernest R. Hilgard
Stanford University

Heinrich Klüver
University of Chicago

Carney Landis
Columbia University

Norman L. Munn
University of Adelaide, Australia

John Paul Nafe
Washington University

Donald G. Paterson
University of Minnesota

Lee Edward Travis
University of Southern California

Robert C. Tryon
University of California

Michael J. Zigler
Wellesley College

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 55

1963

Copyright, 1963, by The Journal Press
Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Anker, James M. 211
Asch, Solomon E. 199
Balow, Irving H. 323
Barry, H., III 145
Bartley, S. Howard 121, 153, 371, 387
Biddle, Bruce J. 183
Bird, Joseph W. 329
Brachman, H. M. 299
Braun, John R. 441
Burton, Arthur 329
Clausen, Johs. 91
Costello, C. G. 299
Craig, Eugene A. 223
Desroches, Harry F. 397
Dittborn, Julio M. 39
England, Samuel J. M. 63
Feldhusen, John F. 357
Fishman, Sidney 345
Ford, Zachary 387
Forney, Robert B. 25
Fuchs, James E. 241
Furth, Hans G. 477
Gaito, John 33
Gates, Paul W. 25
Gengerelli, J. A. 457
Gerjuoy, Herbert 43
Goldstein, Jacob 3
Grosser, George S. 263
Gucker, Donald K. 401
Harleston, Bernard W. 165
Haywood, H. Carl 363
Himelstein, Philip 437
Hughes, Francis W. 25
Jenkins, Thomas N. 49
Jordan, James F. 371
Jordan, Sidney 313
Karrer, Rathe 91
Kew, John K. 287
Kimbrough, Wilson W., Jr. 437
Korin, Hyman 345
Korn, Sam J. 109
Laczek, William J. 263
Lawson, Edwin D. 279
Leaton, R. N. 145
Lehman, P. Joseph 223
Lindner, Marged 199
Littig, Lawrence W. 115
Lordahl, Daniel S. 337
Mayzner, M. S. 469
Morris, James R. 287
Mulholland, Thomas 427
Nelson, Thomas M. 121, 153, 371, 387
Noble, Clyde E. 241
O'Connor, James P. 211
Pasamanick, Benjamin 63
Pearl, David 419
Pressey, Sidney L. 101
Rieber, Robert W. 307
Rogers, Miles S. 445
Ronco, Paul G. 227
Rosensohn, William L. 175
Rosenthal, Robert 33
Roth, Robert Howard 169
Runnals, Sylvia 427
Shoemaker, Donald J. 363
Shure, Gerald H. 445
Soules, Elaine M. 153
Stachnik, Thomas 256
Symmes, D. 145
Thompson, Jack D. 241
Townsend, John C. 211
Tresselt, M. E. 469
Twyman, J. Paschal 183
Walter, Charles L. 397
Warm, Joel S. 287
Weiss, Samuel A. 345
Welch, Livingston 109
Wendt, Miriam 43
Wiener, Charles 3
Yavuz, Halide S. 409
Yeracaris, Constantine A. 115

TABLE OF CONTENTS

On some relations between the perception of depth and of movement . . 3
By Jacob Goldstein and Charles Wiener

Performance in human subjects under delayed auditory feedback after alcohol, a tranquilizer (benzquinamide) or benzquinamide-alcohol combination . 25
By Francis W. Hughes, Robert B. Forney, and Paul W. Gates

The interpretation of levels of significance by psychological researchers . . 33
By Robert Rosenthal and John Gaito

Experimental recollection of dreams 39
By Julio M. Dittborn

Work attitudes of ward attendants for the criminally insane 43
By Miriam Wendt and Herbert Gerjuoy

The primary trait anatomy of the MMPI 49
By Thomas N. Jenkins

The visual appearance of stationary objects: Some theoretical psychophysical implications of the reafference principle 63
By Samuel J. M. England and Benjamin Pasamanick

Comparison of phospene threshold to various visual functions and to electrical sensitivity of the skin 91
By Johs. Clausen and Rathe Karrer

Psycho-technology in higher education versus psychologizing 101
By Sidney L. Pressey

On the elimination of artifacts in GSR conditioning 109
By Sam J. Korn and Livingston Welch

Academic achievement correlates of achievement and affiliation motivations . 115
By Lawrence W. Littig and Constantine A. Yeracaris

Some relations between sensory end results and neural activity in the optic pathway 121
By S. Howard Bartley and Thomas M. Nelson

Familiarization with the test apparatus as a factor in the reinforcing effect of change in illumination 145
By R. N. Leaton, D. Symmes, and H. Barry, III

Visual acuity under conditions of intermittent illumination productive of paradoxical brightness 153
By S. Howard Bartley, Thomas M. Nelson, and Elaine M. Soules

Task difficulty, anxiety level and ability level as factors affecting performance in a verbal learning situation 165
By Bernard W. Harleston

A survey of social psychology courses in teacher training institutions . . . 169
By Robert Howard Roth

A logical method for making a classification of emotions, using Wilhelm Wundt's theory of emotion formation 175
By Willam L. Rosensohn

Role conflict of public school teachers 183
By J. Paschal Twyman and Bruce J. Biddle

A note on "strength of association" 199
By Solomon E. Asch and Marged Lindner

A multivariate analysis of decision making and related measures . . . 211
By James M. Anker, John C. Townsend, and James P. O'Connor

Dimensional analysis in applied psychological research 223
By P. Joseph Lehman and Eugene A. Craig

An experimental quantification of kinesthetic sensation: Extent of arm movement 227
By Paul G. Ronco

A new automatic circuit for the Selective Mathometer (MK II) . . . 241
By Clyde E. Noble, James E. Fuchs, and Jack D. Thompson

Transitional probability in anagram solution in a group setting . . . 259
By Thomas Stachnik

Prior parochial vs. secular secondary education and utterance latencies to taboo words 263
By George S. Grosser and William J. Laczek

Development of patriotism in children—A second look 279
By Edwin D. Lawson

Temporal judgment as a function of nosological classification and experimental method 287
By Joel S. Warm, James R. Morris, and John K. Kew

The relationship between persistence and personality 299
By H. M. Brachman and C. G. Costello

Stuttering and self concept 307
By Robert W. Rieber

The disadvantaged group: A concept applicable to the handicapped . . 313
By Sidney Jordan

Lateral dominance characteristics and reading achievement in the first grade . 323
By Irving H. Balow

Family constellation and schizophrenia 329
By Arthur Burton and Joseph W. Bird

The effect of mediators in concept identification 337
By Daniel S. Lordahl

Pain sensitivity of amputation extremities 345
By Hyman Korin, Samuel A. Weiss, and Sidney Fishman

The effects of small and large group instruction on learning of subject matter, attitudes, and interests 357
By John F. Feldhusen

Measurement of palmar sweating: Effect of repeated measurement from the same area 363
By H. Carl Haywood and Donald J. Shoemaker

Experimental evidence for the involvement of a neurophysiological mechanism in the discrimination of duration 371
By Thomas M. Nelson, S. Howard Bartley, and James F. Jordan

The brightness of part-spectrum targets stimulating the eye parafoveally at intermittency rates below fusion 387
By Thomas M. Nelson, S. Howard Bartley, and Zachary Ford

Daily preferences and institutionalization 397
By Harry F. Desroches and Charles L. Walter

Combining external and internal symbolization in the LSD episode . . . 401
By Donald K. Gucker

The retention of incidentally learned connotative responses 409
By Halide S. Yavuz

Language processing ability of process and reactive schizophrenics . . . 419
By David Pearl

The effect of voluntarily directed attention on successive cortical activation responses 427
By Thomas Mulholland and Sylvia Runnals

A study of self-disclosure in the classroom 437
By Philip Himelstein and Wilson W. Kimbrough, Jr.

Fakability of the Gordon Personal Inventory: Replication and extension . . 441
By John R. Braun

Personality factor stability for three ability levels 445
By Gerald H. Shure and Miles S. Rogers

A method for detecting subgroups in a population and specifying their membership 457
By J. A. Gengerelli

Anagram solution times: A function of word length and letter position variables 469
By M. S. Mayzner and M. E. Tresselt

Classification transfer with disjunctive concepts as a function of verbal training and set 477
By Hans G. Furth

$20.00 per annum
$15.00 per volume
$7.50 per half volume

QUARTERLY
Two volumes per year
Immediate publication

January, 1963
Volume 55, First Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

JANUARY, 1963

(Manuscripts are printed in the order of final acceptance)

On some relations between the perception of depth and of movement . . 3
By Jacob Goldstein and Charles Wiener

Performance in human subjects under delayed auditory feedback after alcohol, a tranquilizer (benzquinamide) or benzquinamide-alcohol combination . 25
By Francis W. Hughes, Robert B. Forney, and Paul W. Gates

The interpretation of levels of significance by psychological researchers . . 33
By Robert Rosenthal and John Gaito

Experimental recollection of dreams 39
By Julio M. Dittborn

Work attitudes of ward attendants for the criminally insane 43
By Miriam Wendt and Herbert Gerjuoy

(OVER)

Copyright, 1963, by The Journal Press
Provincetown, Massachusetts, U. S. A.
Entered as second-class matter January 26, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879
Second-class postage paid at Provincetown, Mass.

The primary trait anatomy of the MMPI 49
By Thomas N. Jenkins

The visual appearance of stationary objects: some theoretical psychophysical implications of the reafference principle 63
By Samuel J. M. England and Benjamin Pasamanick

Comparison of phospene threshold to various visual functions and to electrical sensitivity of the skin 91
By Johs. Clausen and Rathe Karrer

Psycho-technology in higher education versus psychologizing 101
By Sidney L. Pressey

On the elimination of artifacts in GSR conditioning 109
By Sam J. Korn and Livingston Welch

Academic achievement correlates of achievement and affiliation motivations . 115
By Lawrence W. Littig and Constantine A. Yeracaris

Some relations between sensory end results and neural activity in the optic pathway 121
By S. Howard Bartley and Thomas M. Nelson

Familiarization with the test apparatus as a factor in the reinforcing effect of change in illumination 145
By R. N. Leaton, D. Symmes, and H. Barry, III

Visual acuity under conditions of intermittent illumination productive of paradoxical brightness 153
By S. Howard Bartley, Thomas M. Nelson, and Elaine M. Soules

Task difficulty, anxiety level and ability level as factors affecting performance in a verbal learning situation 165
By Bernard W. Harleston

A survey of social psychology courses in teacher training institutions . . . 169
By Robert Howard Roth

A logical method for making a classification of emotions, using Wilhelm Wundt's theory of emotion formation 175
By Willam L. Rosensohn

Role conflict of public school teachers 183
By J. Paschal Twyman and Bruce J. Biddle

A note on "strength of association" 199
By Solomon E. Asch and Marged Lindner

A multivariate analysis of decision making and related measures . . . 211
By James M. Anker, John C. Townsend, and James P. O'Connor

Dimensional analysis in applied psychological research 223
By P. Joseph Lehman and Eugene A. Craig

An experimental quantification of kinesthetic sensation: extent of arm movement 227
By Paul G. Ronco

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 3-23.

ON SOME RELATIONS BETWEEN THE PERCEPTION OF DEPTH AND OF MOVEMENT*

*State University of New York, Downstate Medical Center,
and Board of Education, New York City*

JACOB GOLDSTEIN AND CHARLES WIENER

A. INTRODUCTION

In a paper published in 1958, Gemelli reports that, according to results which he had obtained, left-to-right movement is phenomenally faster than is right-to-left movement when circular disks of equal size are moving simultaneously in opposite directions at equal physical speeds within a rectangular frame. He states that this directional difference held for all of his *S*s and that "none of them (had) the slightest doubt on this point" (9, p. 294). He further specifies that this difference was stronger under conditions of simultaneous than of successive presentation.

On the other hand, unpublished experiments performed by the authors indicate that a right-left differential in phenomenal speed under conditions of simultaneity, even if present, is not as clear-cut and general as Gemelli maintains. These experiments were initially undertaken as a test of Gemelli's hypothesis to the effect that the greater phenomenal speed of left-to-right movement is a function of past experience in reading from left to right. It was originally planned to compare groups of *S*s whose reading experience was primarily from left to right with *S*s whose reading experience was primarily in the opposite direction. Although Gemelli's hypothesis concerning the role of reading experience seemed to be somewhat *ad hoc* in character, such a comparison of two groups appeared to be worthwhile on the ground that positive evidence of group differences would have provided strong support for the empiristic view.[1]

---

* Received in the Editorial Office on August 3, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The preliminary results failed to yield a consistent directional difference in phenomenal speed. However, other findings about to be described induced the authors to pursue the study in a somewhat different direction. Incidentally, Gemelli's finding seems to be in conflict with that of Buytendyk (6), who found that phenomenal speed of lights cast on a screen is greater in the right-to-left than in the left-to-right direction. Interestingly enough, both investigators interpret their results in empiristic terms. Of course, since Gemelli used the apparent movement of black disks while Buytendyk used real motion of a light beam, it is quite possible that the differences in their results were due to differences in the physical characteristics

B. PROCEDURE

The authors' procedure involved the presentation of various stimulus configurations utilizing the simultaneous movements of two disks in opposite directions across a rectangular screen. A 16mm. camera was used in the preparation of the film. A single-frame technique was employed in photographing the successive positions of the disks. Details of the procedure used in the making of the film will be more fully described in a subsequent report.

All of the variations used in the experiments involved the use of circular disks traveling along horizontal pathways parallel to the longer sides of the frame. The distance of the projector from the screen was varied from about five to 15 feet. The images of the disks were therefore of different sizes (3/4-inch to 2 1/2 inches in diameter). The disks were of dull black color so as to minimize glare. The background was yellowish white.

The conditions which were varied included physical speed of movement (the speed of movement of the two disks being always the same within any one presentation but varying from one to another presentation); the extent of the vertical separation between the pathways of the two disks (about eight inches at maximum); and extent of overlap in the time intervals taken by the movements of the two disks (about 1/4 the horizontal distance). In some cases the left-to-right and the right-to-left movements were initiated simultaneously, while in other cases the second disk started on its pathway only after the first disk had moved one fourth of the way.

C. RESULTS

The results do not support the notion of a consistent left-to-right vs. right-to-left difference in phenomenal speed. Under some of the experimental conditions—i.e., where vertical separation was small (about two inches)—the phenomenal difference in speeds between the two disks was barely noticeable by the *S*s and was not consistent with respect to direction. Under some of the other conditions a phenomenal difference in speed was a little more pronounced, but could easily be attributable to factors other than direction (see p. 6). While these preliminary results are by no means sufficient as evidence against the hypothesis that a left-right differential in phenomenal speed does occur, they at least point to the need for more careful specification of the conditions of its occurrence.

---

of their stimuli. Since the published reports of Gemelli's and Buytendyk's studies contain very little information concerning the possibly relevant parametric conditions and experimental controls, it would be difficult to draw conclusions at this stage concerning the factors that might be responsible for the discrepancy.

The results have also yielded data on several aspects of left-right movement which are not discussed in Gemelli's paper:

1. The *S*s' spontaneous comments indicate that the phenomenal speed of disks moving along parallel pathways is not uniform throughout the length of the pathways, but that it becomes noticeably slower when the two disks are in the region of maximum proximity, and again becomes faster as the disks move away from each other on their respective pathways. (The term *region of proximity* will be used to describe the region within which the

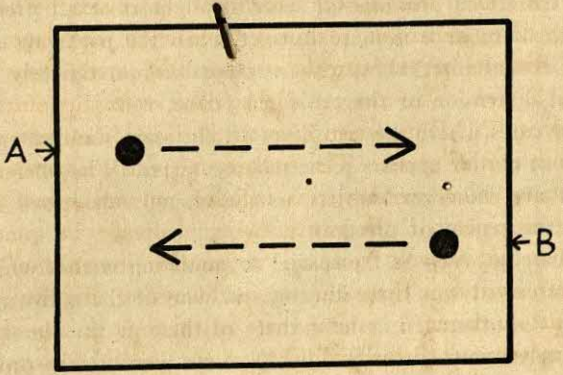


FIGURE 1

BOTH DISKS INTRODUCED SIMULTANEOUSLY

The letters A and B indicate their entrance points on the screen, while the arrows indicate their directions. They move at the same physical speeds along parallel horizontal pathways in opposite directions.

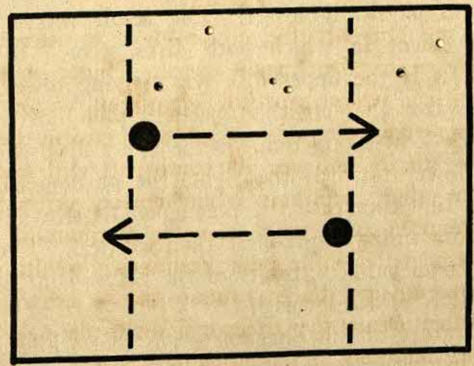


FIGURE 2

THE REGION OF PROXIMITY

In the case of those disk pairs which begin their movement simultaneously this region would coincide with the central area of the rectangle. In terms of a rough estimate the area covered by this region is somewhere between one third and one half of the area of the rectangle.

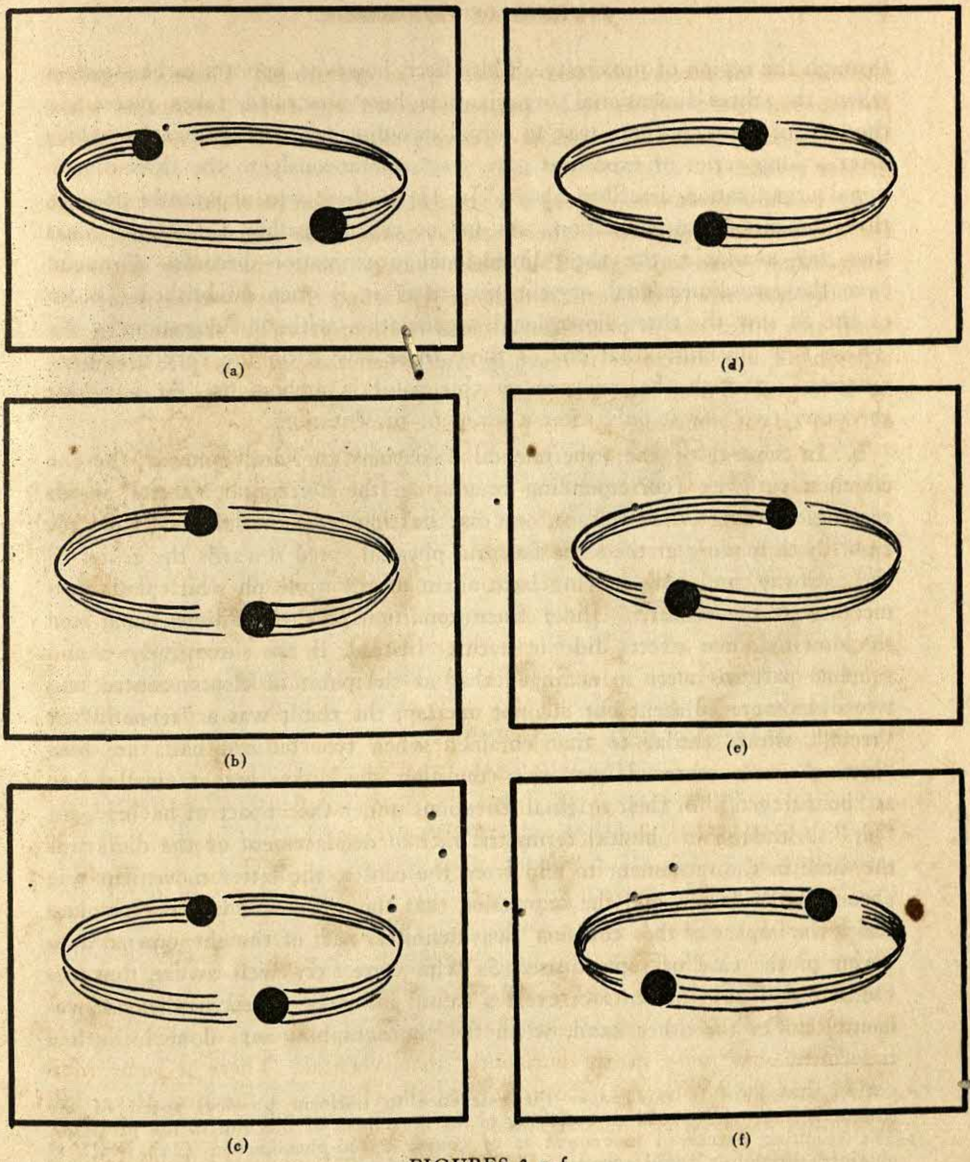
phenomenon of interrelatedness appears to be operating.) These changes in phenomenal speed appear to be particularly striking under conditions when the physical speed of the disks is relatively slow, but are also noticeable at greater speeds.

2. In those cases where there is a vertical separation between the pathways the disk moving along the higher pathway tends to be seen as moving somewhat faster than does the disk which moves along the lower pathway. In those cases where the vertical distance separating the pathways is relatively small this difference in phenomenal speed is slight if at all present. On the other hand, the difference is more distinct when the pathways are separated by a larger vertical interval which corresponds approximately to one third of the vertical dimension of the projected frame.

3. In those cases where the two disks do not start simultaneously, the one which starts out earlier appears to be moving faster. The phenomenal speeds of both disks are, however, subject to the changes described in 1 as they pass through the region of proximity.

4. In the case of seven *S*s the spatial organization within which the movements were perceived was three-dimensional, one of the pathways being seen as closer than the other. In at least three of these *S*s the slowing down and subsequent acceleration (described in 1) were particularly pronounced and were accompanied by some striking changes in the phenomenal pathways through which the disks traveled. These *S*s reported that as the disks approached the region of proximity the disk which was higher appeared to recede further and then again to move nearer, while the reverse held for the disk which was physically lower. The result was a quasi-circular or quasi-elliptoid movement in which both disks seemed to be participating (see Figures 3a-f.). If the upper disk was moving towards the right and the lower disk towards the left this quasi-elliptoid movement was seen as taking place in a clock-wise direction. If the upper disk was moving towards the left and the lower towards the right, the phenomenal movement was counterclockwise. In either case the phenomenal pathway of this elliptoid movement was more nearly horizontal than vertical. There is some indication that the phenomenal eccentricity of the ellipse is a function of vertical separation as well as physical speed.

5. A striking feature of this phenomenon is that while the disks are seen as moving either in the clockwise or in the counter-clockwise direction they are experienced as moving *in relation to each other* rather than independently. Even in the case of *S*s who do not get the three-dimensional effect here described some such feeling of relatedness is present when the two disks pass



FIGURES 3 a-f

THREE-DIMENSIONAL ORGANIZATION

These six figures provide a diagrammatic representation of selected positions in the three-dimensional perceived movement sequence discussed in the text. The curved lines in the drawing were added to indicate something about the direction of the phenomenal movement. The relative sizes of the drawn disks were varied so as to enhance the pictorial representation of three-dimensionality for the reader. It must be emphasized that these drawings are not intended as literal representations of the perceptual experience but are merely to be used as propaedeutic aids in communication.

through the region of proximity. This effect, however, appears to be stronger when the three-dimensional organization here described takes place. In the case of three *S*s there was an initial two-dimensional organization which after a long series of exposures gave way spontaneously to the three-dimensional organization described above. In 15 *S*s there was apparently no such three-dimensional organization. As far as can be gathered from the data thus far available, the three-dimensional organization becomes dominant over the two-dimensional organization after it is once established. Three of the *S*s saw the three-dimensional organization without suggestion by *E*s. There is a possibility that one of these three saw it on the very first presentation, although the evidence on this point is ambiguous. At any rate the other two saw it only after a series of presentations.

6. In several of the experimental variations the disks moved along a common pathway (corresponding roughly to the horizontal bisector of the rectangle serving as the frame), one disk starting from the left, one from the right, both moving at the same uniform physical speed towards the center of the pathway, and both moving back again at the same physical speed after meeting at the center.[2] Under these conditions the three-dimensional and the slowing-down effects did not occur. Instead, if the successively photographed patterns were so arranged that at the point of closest contact the two disks were adjacent but did not overlap, the result was a "rebound" or "recoil" effect, similar to that obtained when two billiard balls are seen "hitting" each other. Under this condition the disks were typically seen as "bouncing off" in their original directions under the impact of having been "hit." Although in physical terms the rate of displacement of the disks was the same in the movement to and from the center, the latter movement was phenomenally faster, and the impression that the movement was taking place under the impact of the "collision" was definitely part of the phenomenal data (even in the case of sophisticated *S*s who were very well aware that the filmed disks could not in fact exert a causal influence on each other's movements).[3] On the other hand, when the photographing was done in such a

---

[2] As elsewhere in this paper the reference to uniform physical speed of the stimuli is to be understood as referring to the movement of the film in the projector. The resulting perceived movement is of course a phi-phenomenon (36) which is obtained through a rapid succession of discrete stimuli.

[3] Some preliminary findings indicate that under conditions of repeated exposure the recoil effect sometimes gives way, at least temporarily, to other perceptual effects. Thus, particularly if the pathways are shortened so that the two filmed projections of the disks are made to move back and forth within the central area of the screen a three-dimensional effect is sometimes obtained whereby the two disks are seen as moving parts of the circumference of a circle (or of a part of a

way as to provide for a stage of partial overlapping, the disks were seen as moving one behind the other and as proceeding on their respective paths without any such rebound effects (and without a change in phenomenal speed). However, the "recoil effect" presents some excellent possibilities for a more detailed study of bidirectional differentials in phenomenal speed. Under conditions just described, just as in the previously mentioned conditions (p. 4ff.) there was no discernible left-right differential in phenomenal speed. If there were any such difference it was completely overshadowed by the "rebound" effect which was bisymmetrical.

D. DISCUSSION

1. *Phenomenal Changes in Speed*

The phenomenal changes in speed which were obtained in the present study are at least superficially in conflict with the findings of Johansson (15), according to whom two objects moving in opposite directions within a rectangular field appear to be moving considerably faster than does a single object moving through a field of the same size at the same physical speed. To be sure, Johansson's experimental situation was not entirely comparable to that of the present study. In Johansson's experiment *S*s were asked to compare phenomenal speeds of objects traveling within two fields. In the present study there was a single field, and *S*s' reports pertained to ***phenomenal changes in speed*** within that field rather than to ***comparisons with phenomenal speed in some other field.***

Although Johansson's finding is thus not really in contradiction to the findings of the present study, an attempt to reconcile the two does create something of a problem. According to Johansson, the greater phenomenal speed produced by the presence of the second moving object represents a special case of the principle of the framework (5, 31), each of the moving objects acting as part of the framework for the other, and thus reducing considerably as well as continuously changing the size of the effective framework (which would otherwise correspond to the entire rectangle). According to Johansson, such a reduction in the effective size of the frame is present even when a stationary object is introduced into the field, but the effect is much stronger when both objects are moving.

---

circle), with the movement seen as being oscillatory, with the two disks or dots maintaining a constant distance from each other on the circumference. While the circle is, of course, "imaginary" in the sense that only the two points on the circumference are seen at any given time, the impression that the movement pathway has a circular shape is nevertheless very strong.

The present evidence indicates that even after prolonged exposure this three-dimensional effect occurs only in relatively few *S*s, and that even in these *S*s the effect is relatively unstable (in the sense that after further exposure it again gives way to the "recoil" effect).

If Johansson is correct in his assumption that two objects moving towards each other within a rectangular frame tend to reduce the effective size of the frame, it would presumably follow that the phenomenal speed of the objects would become greater in the region of proximity. Actually, however, in the present experiment the phenomenal speed of the disks became *slower* rather than faster in this region (see p. 5). This finding would seem to leave open three possibilities:

1. Johansson may have been in error in assuming that his previously mentioned findings could be explained by the principle of the framework.

2. Johansson was correct in his own application of the framework principle to the situation he studied, but the situation of the present study differs from his situation in some important way which must be taken account of before the principle can be correctly applied to it.

3. Johansson was correct in his own application of the framework principle to his own findings, and an application of the same principle would indeed lead to the expectation of acceleration in the region of proximity in the present experiment *if no other factor were operative.* However, there is another factor which acts in the opposite direction, with the results that the expected effect is completely masked.

The first of these three possibilities can be at least tentatively dismissed as unlikely. Johansson's extension of the framework principle seems plausible, and is to some extent supported by data from at least one other source (7). Subsequent research may conceivably show that the differences Johansson obtained in the comparison judgments in the experiment under consideration were due to some factor other than the framework, but for the time being there is no particular reason to doubt his interpretation. Thus, the possibilities are narrowed down largely to Alternatives 2 and 3.

Turning first to Alternative 3, the possibility that in the present experiment the framework effect was counteracted by another quantitatively stronger factor will now be examined in the light of one of Johansson's other findings. According to Johansson, harmonic oscillatory motion of an object along a horizontal pathway gives rise to phenomenal motion which is uniform, except for a brief slowing down at the turning points. This finding, which seems to have been to some extent corroborated by other investigators (11) is open to two possible interpretations. According to the first interpretation, which is favored by Gottsdanker (11) the perception of uniform motion when the stimulus is in oscillatory harmonic motion represents essentially a high degree of "tolerance" of the organism for deviations from the optimal stimulus for the perception of uniform motion. To put this in somewhat different terms, Gottsdanker maintains that there is a fairly wide range of physical oscillatory motions which gives rise to the experience of uniform motion, presumably because uniform perceived motion is for some reason "preferred" over other types of perceived motion. According to an alternative view, which is apparently the view taken by Johansson, though he is not entirely explicit on this point, harmonic oscillatory motion is the optimal stimulus condition for the perception of uniform oscillatory motion. If Gottsdankers' "tolerance" hypothesis is correct, the fact that harmonic oscillatory motion gives rise to perception of uniform motion is of no special relevance in connection with the present findings. On the other hand, if the "optimal stimulus" hypothesis is assumed to be correct, it would provide a possible clue to the nature of the "counteracting" factor which may have had the effect of masking the influence of the framework in the present study. If it is indeed true that harmonic oscillatory motion, which involves considerable physical acceleration in the central area, is the optimal stimulus for perceived uniform oscillatory motion, it would seem to follow that oscillatory motion in which the physical stimulus moves at a uniform rate should give rise to marked phenomenal deceleration in the center. Following this line of reasoning it may be assumed that the slowing down effect obtained in the present study when the two disks were traveling in opposite directions represents a resultant (or some other kind of compromise solution) involving two opposing factors: the slowing down effect which would be expected on the assumption that harmonic motion is the optimal stimulus for the perception of uniform movement, and phenomenal acceleration which would be expected from Johansson's extension of Brown's framework principle (5, 31).

A crucial decision as regards the relative merits of the "tolerance" and the "optimal stimulus" hypothesis cannot be reached without further experiments. However, in view of the importance of this problem, not only for the specific issue under consideration but also for general perceptual theory, some of the relevant evidence will be briefly reviewed.

Quite apart from the fact that such an assumption does not find immediate corroboration in everyday observation of objects moving at physically uniform speeds, it also seems to run counter to some experimental observations. For instance, as Gottsdanker (11) has already pointed out, this assumption does not seem to apply in the case of the Brown experiments on transposition (5), where continuous uniform movement of a series of dots gives rise to phenomenally uniform speed. Although in the original Brown experiments the *S*s were not asked for judgments pertaining to uniformity or lack of uniformity of phenomenal speed, it is unlikely that these *S*s—some of whom were themselves experimental psychologists working on related problems—would have overlooked any marked evidence of phenomenal acceleration or deceleration.

Nevertheless, considerations based upon a comparison with Brown's experiment did not seem to be in themselves quite decisive on this point. In the Brown experiment there is a continuous movement of dots in one direction, and there are thus no definite points which can be distinguished as the beginning or the end of a movement phase. In this respect the Brown experimental situation is very different from the Johansson situation in which movement is oscillatory in character. It might, therefore, be argued that under the conditions of the Brown experiment there is operative some sort of "constancy" process (11, 14) which masks a tendency towards deceleration—and subsequent acceleration—and that such a tendency would perhaps manifest itself under conditions of oscillatory motion.

To test this possibility an attempt has been made to study the phenomenal changes in speed of movement under conditions where a single filmed disk moves back and forth in uniform oscillatory motion along a horizontal pathway across a rectangular screen. The work on this aspect of the problem is still in a preliminary stage, and more elaborate techniques will have to be used before any definite conclusions on this matter can be reached. For instance, in an adequate study of this problem it may be necessary to distinguish between phenomenal acceleration and deceleration as related to the boundaries and the center of: (*a*) the framework within which the movement takes place (5); (*b*) the retinal field (28); and (*c*) the egocentric field (28). However, tentative results do not support the notion of phenomenal changes of the postulated kind in uniform oscillatory movement.

The remaining alternative—namely, that the "transposition" principle is correct but is not directly applicable to the present findings—is supported by the following considerations:

The phenomenal deceleration and subsequent acceleration of the disks occurred under conditions where the pathways of the two disks were geographically parallel but separated by a vertical distance (and it can tentatively be said that these phenomenal changes were optimal only when a critical degree of separation was reached). When the two disks travelled towards each other along a common pathway their speeds appeared to be uniform (see p. 8). To be sure, this result is consistent with the two-factor theory. Since the "framework" effect of the disks on each other is presumably stronger when they are traveling directly against each other on a common pathway, it may be argued that in these instances the effect was sufficiently strong to counterbalance completely those factors which would have favored a slowing-down in the center. There is, however, another consideration which suggests caution before acceptance of this solution. It so happens that, at least within the limits of the study, those stimulus configurations which favor phenomenal deceleration and subsequent acceleration also favor three-dimensional organization. While this may be an irrelevant coincidence, it is at least in line with the possibility—already suggested on other grounds—that the three-dimensionality and the phenomenal changes in speed found in the present study may be intrinsically related. Further experiments will, of course, be necessary in order to determine whether this relationship is in fact an intrinsic one.

A possible relationship between three-dimensionality and alterations in phenomenal speed is to some extent also supported by at least one type of observation based upon a different experimental situation. On prolonged watching of the rotating trapezoidal window (2, 27) one sometimes gets the impression of marked alterations in phenomenal speed, including a slowing-down just prior to the point of phenomenal reversal of direction. In the case of the trapezoidal window—in contrast to the situation of the present experiment—there is, of course, a physical three-dimensionality in the stimulus configuration. However, since in the case of the rotating trapezoidal window there is no consistent correspondence between the physical and the psychological three-dimensionality, the analogy with the present results may be quite relevant. It would seem that in both cases a change in the three-dimensional organization of a movement configuration is accompanied by a change in phenomenal speed (without a corresponding change in the physical configuration).

Thus, a case can be made out in support of the view that three-dimensional organization is the factor that limits the applicability of the framework principle—as defined in terms of two-dimensional retinal geometry—to the present findings. When the three-dimensional effect does occur the phenomenal pathways of the disks are no longer within the same phenomenal plane as the framework, and consequently it may be argued that the framework effect is considerably weakened under this condition when the disks approach the central area of the rectangle. While this assumption would indeed provide a possible explanation of the slowing-down effect for those cases in which the three-dimensional organization does occur, it would leave open the problem of explaining the effect in those cases where such organization does not occur.

One possible way out of this difficulty is to assume that some degree of three-dimensional organization occurs in virtually all cases, but that in some cases it is much stronger than in others, and that in those cases where the three-dimensional organization is not very strong some *S*s are prevented from reporting or consciously registering it by a "stimulus error" (30, 14) which leads them to discard "unrealistic" perceptions. While there is no direct evidence to support this assumption, there is at least one other aspect of the findings which lends it some indirect support, and which makes it at least sufficiently plausible to warrant further study.

As has previously been noted, there is some evidence which indicates that when two disks are traversing the screen at equal speeds in opposite directions along parallel horizontal pathways, the disk which moves along the higher pathway appears to be moving faster than does the disk traversing along the lower pathway.

One plausible way to explain this finding would be to assume that the disk which moves along the higher pathway is at a phenomenally greater distance from the observer, and that—just as in the case of the size-distance relationship (18)—there is a direct relationship between the phenomenal distance of the moving object and phenomenal speed of the object. This explanation would, of course, presuppose the presence of at least a rudimentary three-dimensional organization in all those cases where the movement along the upper pathway is phenomenally faster. On the basis of their own observations of the moving disks the authors are, however, inclined to believe that such a rudimentary three-dimensional organization, if it does exist, is by no means comparable to the fuller three-dimensional organization which apparently emerges only in some *S*s.

Another possibility which may be suggested is that oscillatory motion is an important factor contributing to the formation of the three-dimensional organization.

2. *Stereokinetic Effects*

From a purely descriptive standpoint the three-dimensional effects obtained in the present study can be classified under the heading of stereokinetic phenomena (26, 34), all of which involve three-dimensional effects arising from two-dimensional movement patterns. One of the varieties of stereokinesis is the Kinetic Depth Effect (32) which involves perception of solid objects as a consequence of exposure to changing lines or contours which correspond to the shadows of three-dimensional objects. According to Wallach and O'Connell one condition which is essential for the occurrence of the KDE is the presence of contours or lines which change their length and direction simultaneously. A three-dimensional percept *may* arise if only one of these conditions is fulfilled, as is, for instance, the case with the T figures whose arms undergo a simultaneous and symmetrical change in size without a corresponding change in direction. However, it is Wallach and O'Connell's contention that these two conditions are essential for the universality of the KDE. This position is to some extent modified by Wallach, Weisz, and Adams (34), who add that changing distances and the directions of imaginary lines connecting two points can also be sufficient for producing the KDE.

Another category of stereokinetic phenomena pertains to impressions of depth which arise from two-dimensional patterns in rotation. These phenomena have been investigated by Benussi, by Musatti, and more recently, by Wallach, Weisz, and Adams (see 34), and it is to phenomena of this category that the term "stereokinetic" was originally applied by Musatti. As Wallach, Weisz, and Adams's experiments indicate, some of these phenomena are compelling, and are apparently instances of the KDE, while others—which occur only in some *S*s—are apparently of a more complex nature.

The three-dimensional effects obtained in the present study seem however, to be more closely related to some of the three-dimensional effects described by Johansson (14). Like the latter, they occur only in some *S*s; and also like the latter—they cannot plausibly be explained in terms of a succession of two-dimensional representations of the three-dimensional organization that finally emerges.

According to a theory advanced by Wallach, O'Connell, and Neisser (33), and subsequently to some extent supported by Fishbaine's findings (8), the emergence of three-dimensional forms in the KDE experiment is the result of an accumulation of traces of successive deformations in the two-dimensional representations. In the present study, the three-dimensional effects were obtained under conditions where the distance between the disks and the orientation of the imaginary line connecting them were undergoing constant changes. Using the modified criteria formulated by Wallach, Weisz, and Adams (see above), this effect is indeed a case of KDE. On the other hand, the present findings call into question one of the criteria used by Wallach, Weisz, and Adams (34) in operationally identifying the KDE—namely the criterion of universality. It will be recalled that in the present study less than 20 per cent of the *S*s perceived the three-dimensional pattern without suggestion. Even if those *S*s to whom the three-dimensional pattern was suggested in a "testing-the-limits" procedure are included, the total of those who saw the fullfledged three-dimensional organization amounted to less than half the sample. Thus, although the three-dimensional effects obtained in the present study seem to meet the criteria of the KDE in terms of retinal geometry (and also in terms of phenomenal transformations in those cases where three-dimensional organization does occur), it does not seem to occur with the same degree of regularity as do the KDE effects obtained in the previously cited studies.

While the nature of the initial percepts may very well have been a function of cognitive capacity or of the individual mode of cognitive functioning, it does not by any means follow that the process of transformation to the three-dimensional organization was under cognitive control. The possibility that some sort of "unconscious inference" (13) was involved in this transformation cannot at this stage be completely excluded. However, such introspective reports as are available point in a different direction.

That the three-dimensional effect was a *bonafide* perceptual effect is indicated by the fact that all of the *S*s who saw it described it in terms which were identical in crucial respects. Thus, there was complete agreement as regards the fact that it was the upper pathway that became phenomenally more distant and the lower pathway that became phenomenally closer in the central area. As previously noted, in some cases suggestion was necessary to induce or facilitate the occurrence of the three-dimensional effect. However, even in those cases the *S*s were struck by the suddenness with which this effect emerged. The surprise expressed by these *S*s had indeed some of the quality of the "Aha" experience (19, 21). However, this surprise experience pertained to a change in perceptual organization that had just taken place, if the introspective reports of the small number of *S*s who spontaneously saw the three-dimensional patterns can be taken at face value. The perceptual reorganization itself seems to have had the character of an autochthonous process similar to that involved in the Necker cube reversal rather than that of problem solution based upon insight.

Since in the KDE studies by Wallach *et al.,* the stimulus objects were observed monocularly, while in the present study binocular observation was used, there is the possibility that this difference in procedure accounts for some of the differences in the obtained results. It is, of course, possible that the obtained three-dimensional effects are primarily dependent on monocular factors (10, 25, 32), in which case this difference in the mode of presentation would probably be of minor importance. On the other hand, to the extent that binocular factors are involved, the proportion of *S*s developing the three-dimensional pattern would presumably have been even smaller if only monocular exposures had been used.

It would seem that configurations involving two disks traveling in opposite directions are more ambiguous than are the configurations used in the earlier KDE studies. While the two disks *can* be seen as end-points of an imaginary line undergoing a continuous perspective transformation, they presumably don't *have* to be seen this way. It would also seem that a study of personality and experiential differences between *S*s who readily obtain a three-dimensional organization and those who do not is very much in order (1).

In other words, what is here suggested is that the formation of the KDE is dependent not only on a succession of stimulus patterns, as defined in terms of retinal geometry, but also on the nature of the concomitant percepts, and that these in turn are dependent not only on the retinal configurations but on the total context. While the influence of the total field on perception is, of course, widely recognized in contemporary psychology (e.g., 35, 38, 39), the consequences of this principle for the KDE have not—as far as the authors are aware—been previously spelled out.

3. *The Role of Phenomenal Distance*

The greater phenomenal speed of the disk traveling through the higher pathway is in line with the assumption that the phenomenal distance of the higher pathway is somewhat greater than is the phenomenal distance of the lower pathway—an assumption which corresponds (at least within a wide range) to the situation which, for instance, holds when a photographed landscape is presented on a screen. On the analogy of what is known concerning the size-distance relationship (18) it seems plausible to assume that —other things being equal—the movement which is phenomenally more distant will be also phenomenally faster. Whether or not this explanation is correct can be determined by experiments in which additional factors would be introduced to strengthen or weaken the impression of distance.

However, under conditions where only a "rudimentary" three-dimensional effect is present (see p. 14) the difference in phenomenal distance between the two pathways is presumably relatively slight—an assumption which is in line with the finding to the effect that the reported differences in phenomenal speed are also of relatively small magnitude. When the fullfledged three-dimensional effect develops, the difference in the phenomenal distances of the two pathways becomes, of course, greater—especially in the central area where the elliptoid movement takes place. The changes in phenomenal speed which take place under this condition can, however, no longer be explained merely in terms of the formula whereby the movement which is at a greater distance is phenomenally faster. To be sure, the formula (which is qualified by the phrase "other things being equal") may still be valid in this situation.[4] However, even if valid, this formula can no longer explain the more striking transformations in phenomenal speed which take place under the three-dimensional conditions, i.e., the phenomenal slowing down of the movement (*both* along the lower and the upper pathway) in the central area. This slowing down, it will be noted, occurs in spite of a phenomenal lengthening of the pathway. Paradoxical as this may seem from a common-sense standpoint, this finding exemplifies very strikingly the fact that, given a constant physical speed, along a line approximately perpendicular to the observer's field of view, an increase in the phenomenal length of the pathway may under some conditions be accompanied by a *decrease* in phenomenal speed.

4. *Perception of Causality*

Two of the findings of the present study seem to bear a close relationship to the phenomena studied by Michotte (24), Johansson (14, 15), Heider and Simmel (12), and others under such headings as "perception of causality" and "apparent behavior."

One of these findings pertains to the *feeling of relatedness* (see p. 6) of disks traveling in opposite directions along horizontal pathways. This feeling that the movements of the two disks are dynamically interdependent as they pass through the region of proximity seems to be in the nature of an expressive quality (3, 4) which is presumably autochthonous in character.[5]

The previously mentioned findings concerning the "recoil effect" (see p. 8) parallel closely Michotte's *l'effet lancement* as well as Johansson's (14) findings. A variation which was used by Johansson but which was not included in the present study involved what might be called "vertical overlapping—an arrangement whereby one of the pathways is somewhat higher than the other, but the difference is so slight that there is a partial vertical overlap when the objects are in the same area. (Johansson states that the "recoil" effect is markedly reduced or disappears under this condition.) This seems to be analogous to the finding of the present study concerning the absence of the "recoil" effect under conditions of partial horizontal overlapping.

---

[4] In a recent paper Kilpatrick and Ittelson (17) have presented some evidence which indicates that the principle of size-distance invariance—although approximately correct as an empirical generalization covering a wide range of phenomena —is subject to some striking exceptions, and thus cannot be regarded as an explanatory principle of general validity. Extending the analogy which has previously been suggested (see p. 17) between the role of phenomenal distance in the perception of speed and of size, the further suggestion may be ventured that the inverse relationship between phenomenal speed and phenomenal distance does not hold under all conditions.

[5] The fact that this feeling of relatedness is at a maximum within the region of proximity bears a possible relationship to Wertheimer's law of proximity (37). It should be noted that in the three-dimensional effects here described the phenomenal separation increases as the physical distance along the horizontal pathway decreases. Furthermore, a minimum vertical separation of the pathways seems to be necessary for the three-dimensional organization to occur.

Both in the case of the Johansson study and in the case of the observations here reported the "recoil" effect was obtained under conditions where stimulus arrangements were essentially symmetrical. This point perhaps deserves some emphasis because in this respect both types of arrangements are different from those typically used by Michotte (24).

5. *Some Relations to the Gestalt Laws of Organization*

Several authors (e.g., 16, 22, 23 and especially 14) have taken the view that the laws of organization pertaining to movement are not identical with the laws of organization applicable to stationary configurations.[6] While this is not the place for a detailed discussion of this issue, the fact is that the principle of proximity as applied in the preceding comments is not identical with the law of proximity applied by Wertheimer to stationary configurations. For instance, the fact that under some conditions a minimum of vertical separation appears to favor the emergence of a feeling of relatedness (see pp. 6-8) does not seem to be derivable from the available knowledge concerning stationary groupings.

Without attempting a systematic analysis of the similarities and differences between the principles of organization applicable to static configurations and those applicable to configurations involving movement, it is, nevertheless, possible to note that the stereokinetic effects obtained in the study seem to be in line with the principles of *closure,* of *symmetry,* and of *good continuation.*

The operation of *closure* seems to be exemplified by the fact that when the depth effect occurs the two pathways are experienced as forming a phenomenally elliptoid figure, even though physically the two pathways run along parallel lines, and even though the vertical distance separating the two pathways amounts to as much as one third of the vertical dimension of the frame.

There is also some evidence of the operation of a principle of *symmetry,* since the lower pathway is curved in the direction roughly opposite from that of the upper pathway, and since both in the case of the upper and of the lower pathway the elliptoid shape is roughly symmetrical as between the left and the right sides. (There is, of course, no implication here that the elliptoid shape is actually symmetrical, except in a crudely topological sense.)[7]

---

[6] For a recent discussion of the Gestalt laws of organization see Rock (29).

The principle of *good continuation* also seems to be operative in that the distances between the phenomenal pathways change gradually and continuously rather than by discrete steps.

As far as the factor of *similarity* is concerned, it seems plausible to conjecture that the stereokinetic effects obtained in the study would have been much more difficult to produce if, for instance, the two disks had been of very different size or shape or if they traveled at different physical speeds.[8]

6. *Bearing on Gemelli's Study*

In contrast to the findings reported by Gemelli, the present findings do not support the notion that left-to-right visual movement is phenomenally faster than is simultaneous right-to-left movement when the physical speeds are equal. The possibility that a genuine left-right differential in phenomenal speed does exist is indeed left open by the results obtained thus far, but the results indicate that such an effect, if at all present, is not as clear-cut and general as reported by Gemelli. Moreover, the present results suggest that factors other than direction of movement may affect the relative phenomenal speed of disks moving with equal physical velocity in opposite directions. For instance, there is some indication (see p. 6) that the disk which is traveling through the pathway which is located higher on the screen is seen as moving faster. Since Gemelli's paper contains no information as to whether such additional factors were controlled in his study, the possibility remains that his findings are to some extent a function of this factor or of some other similar factor.

It may further be noted that Gemelli makes no mention of phenomenal changes in speed or direction as the disks traveled across the screen. This may mean that his *S*s did not experience such changes or that such changes were neglected because the procedure which he used called only for comparison judgments. In addition there is also the possibility that the oscillatory motion which was used in the present experiments but apparently not in Gemelli's was the crucial factor in producing the three-dimensional organization and the correlated transformations in phenomenal speed.[9] In view of the present findings, which indicate that phenomenal changes in speed take place as two disks approach each other—and, less directly, in view of Johansson's contention that uniform phenomenal speed is produced by harmonic oscillatory motion (see pp. 10ff.)—the question arises as to whether phenomenal changes in speed, if they did take place, may have affected Gemelli's data. Whether a left-right differential in phenomenal speed does exist can only be determined by a future study. What the present results do indicate quite clearly is that such a study—if it is to be valid—must take into account factors which may be of crucial importance but which were apparently uncontrolled by Gemelli.[10]

---

[7] In a subsequent paper the authors will report in more detail on some of the conditions which favor the emergence of circular as distinguished from elliptoid phenomenal movement pathways.

[8] From a descriptive standpoint the phenomena of three-dimensional organization reported in this paper may be considered as "satiation" phenomena, at least in the sense that they involve perceptual changes apparently resulting from repeated exposures. It would seem that these phenomena bear a somewhat closer resemblance to the trapezoidal window oscillatory phenomena than they do to the "satiation" phenomena described by Köhler and Wallach (20).

E. SUMMARY

Gemelli's findings concerning directional differences in phenomenal speed were subjected to an experimental test. No evidence of consistent right-left differences in phenomenal speed was found. Other effects were obtained, none of which are described in Gemelli's paper. These included phenomenal changes in speed and phenomenal depth effects under conditions where two filmed disks were moving in opposite directions at the same uniform physical speeds along parallel horizontal pathways. Under other conditions "recoil" effects were obtained.

The findings are discussed in relation to the Gestalt laws of organization, the framework as a determinant of phenomenal speed, stereokinetic phenomena (including the KDE), phenomenal distance, and in relation to various studies in the field of perception.

REFERENCES

1. ALLPORT, G. W., & PETTIGREW, T. F. Cultural influences on the perception of movement: the trapezoidal illusion among Zulus. *J. Abnorm. & Soc. Psychol.*, 1957, **55**, 104-113.
2. AMES, A., JR. Visual perception and the rotating trapezoidal window. *Psychol. Monogr.*, 1951, **65**, No. 7 (Whole No. 324).
3. ARNHEIM, R. The Gestalt theory of expression. *Psychol. Rev.*, 1949, **56**, 156-171.
4. ———. Art and Visual Perception. Berkeley and Los Angeles: Univ. Calif. Pr., 1954.
5. BROWN, J. F. The visual perception of velocity. *Psychol. Forsch.*, 1931, **14**, 199-232.

---

[9] Additional experiments intended to test this possibility are under way.

[10] In research currently under way the authors are investigating in more detail some of the parametric conditions related to the optimal occurrence of the phenomena here described. Some of this work involves the use of a forced-choice technique in which the *S* is asked to select one of a pair of visual patterns as matching more closely his movement experience. The role of past experience in inducing the three-dimensional effects is also under study.

6. Buytendyk, F. Les différences sensorielles et motrices entre la gauche et la droite. *Acta Psychol.*, 1959, **15**, 276-277.
7. Duncker, K. Über induzierte Bewegung. *Psychol. Forsch.*, 1929, **12**, 180-259.
8. Fishbaine, E. D. The Role of Memory Traces in the Perception of the Kinetic Depth Effect. Unpublished doctor's dissertation, Graduate Faculty of Political and Social Science, New School for Social Research, 1954.
9. Gemelli, A. The visual perception of movement. *Amer. J. Psychol.*, 1958, **71**, 291-297.
10. Gibson, J. J. The visual perception of objective motion and subjective movement. *Psychol. Rev.*, 1954, **61**, 304-314.
11. Gottsdanker, R. M. The ability of human operators to detect acceleration of target motion. *Psychol. Bull.*, 1956, **53**, 477-487.
12. Heider, F., & Simmel, E. A study of apparent behavior. *Amer. J. Psychol.*, 1944, **57**, 243-259.
13. Helmholtz, H. von. Physiological optics. (*Ed.*, J. P. C. Southall). *Optical Soc. Amer.*, 1925, Vol. **3**.
14. Johansson, G. Configurations in Event Perception—An Experimental Study. Uppsala: Almqvist & Wiksells, Boktryckeri, 1950.
15. ———. Configurations in the perception of velocity. *Acta Psychol.*, 1950, **7**, 25-79.
16. Katz, D. Gestalt Psychology (Trans. by R. Tyson). New York: Ronald, 1950.
17. Kilpatrick, F. P., & Ittelson, W. H. The size-distance invariance hypothesis. *Psychol. Rev.*, 1953, **60**, 223-231.
18. Koffka, K. Principles of Gestalt Psychology. New York: Harcourt Brace, 1935.
19. Köhler, W. Gestalt Psychology (2nd ed.). New York: Liveright, 1947.
20. Köhler, W., & Wallach, H. Figural after-effects: An investigation of visual processes. *Proc. Amer. Philos. Soc.*, 1944, **88**, 269-357.
21. Kris, E. On preconscious mental processes. *Psychoanalytic Quart.*, 1950, **19**, 540-560.
22. Metzger, W. Beobachtungen über phänomenale Identität. *Psychol. Forsch.*, 1934, **19**, 1-60.
23. ———. Tiefenerscheinungen in optischen Bewegungsfeldern. *Psychol. Forsch.*, 1935, **20**, 195-260.
24. Michotte, A. La Perception de la Causalité (2nd ed.). Louvain: l'Institut supérieur de Phil., 1954.
25. Motokawa, K., Nakagawa, D., & Kohata, T. Monocular stereoscopic vision and gradients of retinal induction. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 392-403.
26. Musatti, C. L. Sui fenomerri stereocinetici. *Arch. Ital. di Psicol.*, 1924, **3**, 105-120.
27. Pastore, N. Some remarks on the Ames oscillatory effect. *Psychol. Rev.*, 1952, **59**, 319-323.
28. Rock, I. The perception of the egocentric orientation of a line. *J. Exp. Psychol.*, 1954, **48**, 367-374.
29. ———. The present status of Gestalt psychology. *In* Peatman, J. G., & Hartley, E. L. (*Eds.*), *Festschrift for Gardner Murphy*. New York: Harper, 1960. Pp. 117-144.
30. Titchener, E. B. A Textbook of Psychology. New York: Macmillan, 1909-1910.
31. Wallach, H. On the constancy of visual speed. *Psychol. Rev.*, 1939, **46**, 541-552.
32. Wallach, H., & O'Connell, D. M. The kinetic depth effect. *J. Exp. Psychol.*, 1953, **45**, 205-217.

33. WALLACH, H., O'CONNELL, D. M., & NEISSER, U. The memory effect of visual perception of third-dimensional form. *J. Exp. Psychol.*, **45**, 360-366.

34. WALLACH, H., WEISZ, A., & ADAMS, P. A. Circles and derived figures in rotation. *Amer. J. Psychol.*, 1956, **69**, 48-59.

35. WERNER, H., & WAPNER, S. Toward a general theory of perception. *Psychol. Rev.*, 1952, **59**, 324-338.

36. WERTHEIMER, M. Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.*, 1912, **61**, 161-265.

37. ———. Untersuchungen zur Lehre von der Gestalt II. *Psychol. Forsch.*, 1923, **4**, 301-350.

38. WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. Psychological Differentiation: Studies of Development. New York: Wiley, 1962.

39. WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. Personality through Perception. New York: Harper, 1954.

Psychopharmacology Research Unit
State University of New York
Down-State Medical Center
450 Clarkson Avenue
Brooklyn 3, New York

804 Lenox Road
Brooklyn 3, New York

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 25-32.

PERFORMANCE IN HUMAN SUBJECTS UNDER DELAYED AUDITORY FEEDBACK AFTER ALCOHOL, A TRANQUILIZER (BENZQUINAMIDE) OR BENZQUINAMIDE-ALCOHOL COMBINATION*

Department of Pharmacology, Indiana University School of Medicine

FRANCIS W. HUGHES, ROBERT B. FORNEY, AND PAUL W. GATES[1]

A. INTRODUCTION

A delayed auditory return of an *S*'s voice results in a defect in speech performance. This is manifested by repetitions of words and phrases, slurring, and stuttering (1). The period in which this occurs has been described by *S*s as times of surprise and frustration (11, 12). From observations of *S*s under such conditions, we have believed that this system might be utilized for the self induction of an "anxiety-state." Performance under this self-induced stress can be quantitated (5) and the effectiveness of agents such as sedatives, tranquilizers or alcohol can be evaluated for their capacity to reduce or enhance performance. The data presented in this paper represent such an evaluation using alcohol in low concentrations and a short acting tranquilizer (benzquinamide) as well as combinations of these drugs. A placebo drug and placebo alcohol were also included.

B. METHOD

1. *Subjects*

Sixteen paid volunteer medical and graduate students were included in the study, eight males and eight females. Ages ranged from 20 to 48 with a median of 24. The median weight was about 148 pounds, ranging from 96 to 220. *S*s were encouraged to perform well; in fact to compete with their fellow students.

2. *Apparatus*

S spoke into a Sure 55S microphone which was fed into an Ampex 350 recorder operating at 7-1/2 inches per second. A playback head was mounted

---

* Received in the Editorial Office on August 7, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported (in part) by a PHS research grant AC-20 from the Division of Accident Prevention, Bureau of State Services, Public Health Service.

to bypass the recording system and return a delayed audio feedback into ear phones with rubber caps. *S* sat with his face six inches from the microphone and was instructed to read and speak naturally. The intensity of sound was maintained higher than a comfortable listening level so that *S* could not ignore it. A soundproof carpeted room was used in which only *S* and technician were present.

3. *Tests*

The tests were modifications of those previously described (5). Nine tests were used and each test was timed and recorded for two minutes. Different variations of the same test were used in order to prevent familiarization with the reading material or problems. A short unrecorded practice session was allowed each *S* at the beginning of the test period. The eight tests were as follows:

Test 1. *Verbal Output*. *S* was instructed to read as rapidly but distinctly as possible for two minutes, a passage from "Aristotles Selections." The material that was read was not completely comprehensible to *S* in that it was necessary to have read the previous pages in order to fully understand the meaning of the text.

Test 2. *Reverse Reading*. *S* read a similar passage but in reverse, i.e., the last word on the page was read first and reading was from right to left and upward.

Test 3. *Reverse Count*. *S* was instructed to start at a number such as 10,000, 500 or 100 and count in reverse rapidly.

Test 4. *Progressive Count*. *S* was to start at a number and add a specified number continually for two minutes, e.g., 1, 4, 7, 10, 13 etc. or 6, 11, 16, 21 etc.

Test 5. *Addition*. Typed simple addition problems were given to *S* and he was to read the problem and provide the final answer enunciating each word, e.g.,

Seven plus twelve equals ____________
Nineteen plus three equals ____________

Test 6. *Subtraction*. Similar to test 5 in format but using simple problems in subtraction.

Test 7. *Addition plus seven*. *S* was to say every word but mentally add seven to the final answer. Using example in test five above the subject would *say*

Seven plus twelve equals twenty-six

Nineteen plus three equals twenty-nine

S was not to say the seven aloud but quietly include it in the final answer. An initial screening of other *S*s indicated that seven was the most difficult of the numbers under ten to so include in these problems.

Test 8. *Subtraction plus seven.* Similar to test 7 except that a series of problems in subtraction were presented to *S,* and he was asked to non verbally add seven to the final answer.

Test 9. *Color Test.* This was a modification of the Stroop test (14). A sheet of paper had printed on it names of colors (red, yellow, blue, green, orange, violet, brown, black). The names were printed in different colors. *S* was instructed to read the name of the first word and the color of the second printed word and continue in this alternate pattern for two minutes.

4. *Scoring*

The tests were evaluated by a non-participating technician. Errors were enumerated as per evaluation sheet reproduced in previous publication (5). Output per error was calculated for each test. This comprised the number of words or numbers completed divided by the number of errors in the same two-minute period. Although *S* might continue beyond the allotted time, only two minutes of each recorded test was graded.

5. *Experimental Drugs*

Benzquinamide (Quantril) is a new azabenzoquinolizine compound which disrupts non-discriminated avoidance behavior in rats; it only weakly depletes norepinephrine and serotonin concentrations in rat brains in high dosages (13, 15). It was selected as a drug in this study because preliminary reports of clinical studies indicate that a minimal sedative action is elicited at an effective human tranquilizing dose (4). *S*s were given 25 mg. of the drug in capsule form.

Alcohol has been studied by these and other authors for its capacity to reduce performance in humans. Most of these studies have been at a dosage level which effects blood levels of 0.10 per cent or greater. In the present study, alcohol was administered in a dosage level to effect a blood level of approximately 0.05 per cent (50 mg. per cent). *S*s drank 100 proof bonded bourbon whiskey in quantities calculated to contain 45 ml. of 200 proof alcohol per 150 lb. body weight. This was ingested in seven-up and ice.

Placebo drugs were administered as well as placebo bourbon. The latter was prepared by aerating seven-up and adding 2 ml. of artificial bourbon flavoring (Fritzsche Brothers, Inc., N.Y., N.Y.). These drinks were prepared by a technician and given to *S* already mixed. *S* was instructed to drink the beverage over a 30-minute period. The capsule was taken at the onset of drinking. One-half hour after completion of the drink his blood alcohol level was determined from an alveolar air analysis using the Borkenstein Breathalyzer (2). Simulated analyses were also performed when *S* had the placebo alcohol. Each *S* performed the series of tests on four occasions. He had the following four treatments: (*a*) benzquinamide plus alcohol; (*b*) placebo drug plus alcohol; (*c*) benzquinamide plus placebo alcohol; and (*d*) placebo drug plus placebo alcohol.

A 4 × 4 Latin square design was used.

At the completion of the test another blood alcohol was determined and the mean calculated from the concentration at beginning and end of the test was recorded. At this time he completed a Cornell Medical Index Health Questionnaire (3) and also indicated whether or not he believed that he had a drug or placebo, alcohol or not, and estimated the type of drug (stimulant or depressant) and also estimated his blood level of alcohol. The students had received lectures on the effects of alcohol at various concentrations and so they were more qualified than the usual *S* to make such a prediction.

C. Results

A statistical evaluation is shown in Table 1 for a significant *F* value of 2.815 at 0.05 confidence level. An individual comparison among means was made; any two means that differ by more than the value obtained from q.05 ($K = 4$, $df = 45$) $\overline{\mathrm{SY}}$ is not considered to be significantly important. It should be noted that the *F* value is different for Tests 7, 8 and 9. This is because the number of *S*s was eight rather than sixteen. Three new tests were added to the protocol of the last eight *S*s of the study. Tests 3, 4 and 9 show no significant difference at this level. In all other tests *S*s' efficiency was greater after the drug than after alcohol. In all tests the efficiency score (output/error) increased from control figures after the drug and diminished after alcohol was administered. Although there was a trend of improvement with drug and a deficiency with alcohol in all cases, there was no apparent antagonism of alcohol by benzquinamide. This is not entirely surprising and one might expect rather synergistic action if one believes that alcohol and the tranquilizers should be generalized as depressant drugs.

A simulated alcohol is a difficult placebo to prepare. If any *S* is duped,

TABLE 1
MEAN OUTPUT/ERROR FOR TESTS AND ANALYSIS OF VARIANCE

Test #	F-value 0.05 level = 2.815	Minimum difference for sig. at .05 level	Mean output/error			
			No drug no alc.	No drug alc.	Drug no alc.	Drug alc.
1	4.243	6.490	14.993	11.158	19.298	12.699
2	5.670	9.621	13.327	9.504	22.656	9.349
3	0.101	5.139	3.090	3.545	3.248	2.511
4	2.666	3.415	5.577	3.850	5.917	5.116
5	11.408	4.339	12.300	5.234	13.390	7.485
6	5.407	5.256	10.671	4.955	12.493	7.155
7*	6.184	4.669	5.593	4.005	9.791	3.079
8*	6.211	0.550	5.948	3.089	7.771	3.256
9*	1.272	5.284	9.333	7.864	9.480	6.244

* F-value for 0.05 confidence level is 3.070 because of only eight subjects. Source of variation is the same in each case namely; subjects, treatment and error. For Tests 7, 8 and 9 the comparison among the means was obtained from q.05 (k = 4, df = 21) $\overline{SY}$. Degrees of freedom for the first six tests are fifteen, for Tests 7, 8 and 9 there are seven.

TABLE 2
IMPRESSIONS OF SUBJECTS CONCERNING ALCOHOLIC CONTENT OF DRINKS AND TYPE OF DRUGS

Actual treatment	No drug no alc.	No drug alc.	Drug no alc.	Drug alc.
Nos. believing had alcoholic beverage	3	14	1	15
Nos. believing had drug	2	6	1	6
Believe it was stimulant drug	1	2	0	3
Believe it was depressant drug	1	4	1	3
Drug added on to effect of alcohol	0	6	0	3

the placebo effect is partially successful. In this study (Table 2) three of sixteen *S*s with no drug and one of sixteen with drug thought they had received an intoxicating beverage although they had a simulated alcoholic beverage. As would be expected most of the *S*s who received alcohol, knew this by the end of the experiment. On the other hand, a few of the *S*s receiving alcohol were convinced that they had not. From the same table it is apparent that it is easy to control studies with drugs by using placebos. The general pharmacologic activity of the drug in question was not predictable by the *S*s.

Six of the people treated with placebo drug and alcohol, believed that they had a drug and in fact that it added on to the general depression produced by alcohol. Of those receiving both alcohol and drug six believed they had a drug but only three of these thought it added on to alcohol. The opinions of *S*s concerning the drug had no correlation with actualities of the experimental design.

Data from the Medical Index Questionnaire (Table 3) indicate that no difference in symptoms was evident. Our findings indicate that benzquinamide

TABLE 3
SYMPTOM QUESTIONNAIRE

No. of symptoms checked	No drug no alc.	No drug alc.	Drug no alc.	Drug alc.
0-4	14	9	12	7
5-9	2	1	3	1
10-14	0	2	1	7
15-19	0	1	0	1
20-24	0	3	0	0

is a tranquilizer which causes no gross subjective depression either alone or in combination with low levels of alcohol.

D. Discussion

Most people can experience a generalized depressant response with low levels of alcohol. Under special conditions such as social gatherings these same people may manifest exhilaration. This apparent excitation may actually result from a depression of specific areas of the central nervous system (10). Most studies have shown that mental and physical performance is impaired although social performability may be improved. Because of this depression, a subjective evaluation of a concomitantly administered drug may be a confused one.

In a previous study we have shown that automobile driving performance is impaired with blood levels of alcohol as low as 50 mgm per cent (6). Any drug which may add on to the depression due to alcohol furthers the deficiencies induced by alcohol in the body. The potentiation of effects of alcohol by tranquilizers is relatively easy to demonstrate in smaller animals (7, 8, 9). A practical application of the findings in this study, as well as previous ones, may pertain to skilled performance of members of our population who are taking therapeutic doses of tranquilizers and social quantities of ethanol.

The measurement of performance in this study, however far removed from driving, illustrates that a tranquilizer (benzquinamide) improved performance and alcohol decreased it. It also adds to the contention of the authors that alcohol is not a tranquilizer in any "dosage" or in any form (7). A tranquilizer theoretically should not depress or sedate a patient and should not impair performance. In fact removal of anxiety stimuli or induction of nonreactivity to it may result in improvement of performance.

The performance of *S*s was measured while they were under the stress of a self-induced anxiety situation. This procedure may have advantages in that a potential synergistic action of alcohol and a tranquilizer may be explored in a laboratory situation before extension of the studies to more complicated but natural activities. It appears that this apparatus may be useful in the evaluation of a new drug for possible tranquilizing action.

E. Summary

The present study investigated the verbal performance of subjects under a stimulus of delayed auditory feedback after receiving a tranquilizer (benzquinamide), alcohol or benzquinamide-alcohol combinations.

It was found that: (*a*) benzquinamide increased performance while under

this self-induced anxiety stimulus; (*b*) alcohol effected a deficiency in performances in low concentrations 50 mgm per cent; and (*c*) synergism of this tranquilizer and alcohol was not evident.

REFERENCES

1. BLACK, J. W. The effect of delayed side-tone upon vocal rate and intensity. *J. Speech Hearing Dis.,* 1951, **16**, 56-60.
2. BORKENSTEIN, R. F., & SMITH, H. W. The breathalyzer and its applications. *Med. Sci. Law,* 1961, **2**, 13-22.
3. CORNELL MEDICAL INDEX HEALTH QUESTIONNAIRE. New York: Cornell Univ. Medical College, 1945.
4. FELDMAN, P. E. Clinical evaluation of benzquinamide (P-2647) in the control of tension states and "hyper-syndromes." *Psychosomatics,* 1962, **3**, 148.
5. FORNEY, R. B., & HUGHES, F. W. Delayed auditory feedback and ethanol: Effect on verbal and arithmetical performance. *J. of Psychol.,* 1961, **52**, 185-192.
6. FORNEY, R. B., HUGHES, F. W., HULPIEU, H. R., & DAVIS, C. A. Performance in a Gymkhana sports car event with low levels of blood alcohol. *Traffic Safety Res. Rev.,* 1961, **5**, 8-12.
7. FORNEY, R. B., & HUGHES, F. W. Behavioral effects on the rat of benzquinamide and benzquinamide-alcohol combinations. *Arch. Int. Pharmacodyn.* (in press).
8. HUGHES, F. W., & ROUNTREE, C. B. Influence of alcohol-tranquilizer combinations on choice-discrimination in rats. *Arch. Int. Pharmacodyn,* 1961, **133**, 418-432.
9. HUGHES, F. W., ROUNTREE, C. B., & FORNEY, R. B. Suppression of learned avoidance and discriminative responses on the rat by chlordiazepoxide (librium) and ethanol-chlordiazepoxide combinations. *J. Genet. Psychol.* (in press).
10. HULPIEU, H. R., & HARGER, R. N. The alcohols. *In Pharmacology in Medicine,* V. A. Drill (*Ed.*). New York: McGraw-Hill, 1958.
11. LEE, B. S. Some effects of side-tone delay. *J. Acoust. Soc. Amer.,* 1950, **22**, 639-640.
12. ————. Effects of delayed speech feedbacks. *J. Acoust. Soc. Amer.,* 1950, **22**, 824-826.
13. LOMBARDINO, J. G., BODEN, J. I., GERBER, C. F., MCLAMORE, W. M., & LAUBACK, G. D. Azabenzoquinolizine with tranquilizing activity. *J. Med. & Pharm. Chem.,* 1961, **3**, 505-517.
14. STROOP, J. R. Studies of interference in serial verbal reactions. *J. Exper. Psychol.,* 1935, **18**, 643-662.
15. WEISSMAN, A., & FINGER, K. F. Effect of a new benzoquinolizine on avoidance behavior and brain amine levels. *Fed. Proc.,* 1962, **21**, 420.

Indiana University Medical Center
Indianapolis, Indiana

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 33-38.

THE INTERPRETATION OF LEVELS OF SIGNIFICANCE BY PSYCHOLOGICAL RESEARCHERS*[1]

University of North Dakota and Kansas State University

ROBERT ROSENTHAL AND JOHN GAITO

A. PURPOSE

The interpretation of a "level of significance," or probability of making an error of the first kind, has a number of important implications. The obtained level of probability (*p*) may be interpreted by one investigator as low enough to establish a real difference but not by another. The decision to publish research findings may hinge on the interpretation of *p* levels, where the decision may be that of the investigator and/or that of an editor. A number of workers have discussed how one should or might interpret a level of significance (McNemar, 1955; Ray, 1960; Siegel, 1956; Snedecor, 1956). To our knowledge, however, no data have been available bearing on the question of how psychological researchers actually do interpret various levels of significance. The purpose of the present study was to obtain preliminary data bearing on this question.

B. METHOD

1. *Subjects*

Nine members of the UND Psychology Faculty, all holding the Ph.D., and 10 graduate students in the same department comprised the sample. All 19 *S*s were actively engaged in empirical, quantitative research at the time of the present study.

2. *Procedure*

*S*s completed a questionnaire requesting them to rate their degree of belief or confidence in a variety of *p* levels: once assuming an *n* of 10, and then assuming an *n* of 100. The possible ratings they were to assign to each *p* level were as follows:

5 extreme confidence or belief
4 great confidence or belief
3 moderate confidence or belief

---

* Received in the Editorial Office on August 7, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Preparation of this paper was facilitated by a research grant (G 17685) to R.R. from the Division of Social Sciences of the National Science Foundation.

2 mild confidence or belief
1 minimal confidence or belief
0 complete absence of confidence or belief

C. RESULTS

Fourteen *p* levels were rated; however, all *S*s rated the .70 and .90 levels as zero for both conditions of sample size and these were excluded from further analysis. The analysis of variance of the ratings is shown in Table 1. The Groups × P Levels × Sample Size (G × P × S) interaction was used as the error for testing both Groups × P Levels (G × P) and Sample Size × Groups (S × G) (assuming that σ^2_{GPS} equals zero). All *S*s interaction components were also tested by G × P × S. Sample Size (S) was tested by *S*s × S Within Groups; P Levels (P), by *S*s × P Within Groups; Sample Size × P Levels (S × P), by *S*s × S × P Within Groups; and Groups (G), by *S*s Within Groups.

TABLE 1
ANALYSIS OF VARIANCE OF RATINGS

Source	*df*	Mean square	*F*
Between *S*s	18		
Between Groups (G)	1	257.80	4.21
Between *S*s Within Groups	17	61.27	
Within *S*s	437		
Sample Size (S)	1	96.70	22.75*
*S*s × S Within Groups	17	4.25	6.25*
P Levels (P)	11	893.35	13.40*
*S*s × P Within Groups	187	66.65	98.01*
S × P	11	2.04	< 1
*S*s × S × P Within Groups	187	9.94	14.62*
S × G	1	.80	1.18
G × P	11	9.10	13.39*
G × P × S	11	.68	

* Significant at or below .01 *p* level.

The G × P interaction source was significant. This indicated that differences in ratings between faculty and graduate students were not the same over all *p* levels. In a situation such as this, tests of the main effects can be meaningless (Gaito, in press). Thus the G × P interaction was graphed to note the type of interaction (Figure 1). The plot showed that the ratings of the graduate students were higher arithmetically at every *p* level than were the ratings of the faculty. Because the two curves did not overlap or cross at any point, tests of main effects are meaningful. Such tests are unnecessary, however, in that the G × P interaction has provided the basic information required. Yet tests of these main effects are shown in Table 1 for completeness.

Sample Size, P Levels, and all interactions containing *S*s were significant. The *p* value for the *F* test of Between Groups was .06. However, the G × P interaction had already indicated that the two groups' differences vary over *p* levels. When interactions are present, *F* tests of main effects (which involve average effects) can sometimes provide misleading inferences.

The G × P × S interaction is plotted in Figure 1. The G × P interaction is also included in the plot. The differences present in the graph are due to Group, P Level, S, and G × P differences. G × P × S can not be tested to determine if σ^2_{GPS} is greater than zero.

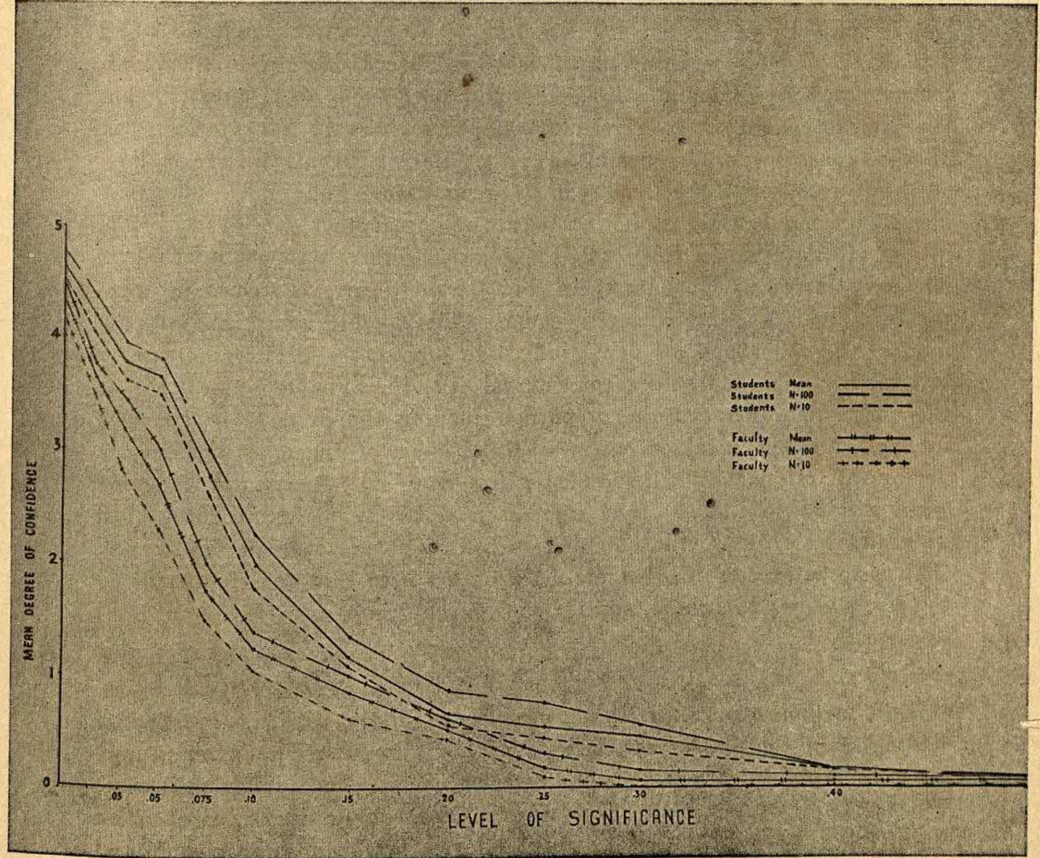


FIGURE 1

The plot of the Groups × *P* levels × Sample Size interaction. The Groups × *P* levels interaction (which was significant) consists of the curve for graduate students and the one for faculty, each averaged over both sample sizes.

Figure 1 shows clearly not only that greater confidence was placed in lower *p* levels but also that greater confidence (1/3 of a unit of confidence, on the average) was placed in *p* levels based on the larger *n*. In addition, graduate students tended to show greater confidence (1/2 of a unit of confidence, on the average) in any given *p* level than did faculty members. Both the curves for faculty and graduate students *S*s appeared to be of exponential nature.

The variability of ratings was greater at lower *p* levels, the standard deviation ranging from .99 to .74 for *p* levels between .001 and .20. At .25, .30, .40, and .50 levels, the standard deviations were .64, .52, .32, and .24, respectively.

Table 2 shows the proportion of *S*s who judge a decrease in their confidence across a particular interval of significance levels to be greater than, less than, or equal to their decrease in confidence across the following interval of significance levels. In moving over the interval .01–.03 we are adding, on the average, two Type I errors in 100 replications of our study. Similarly, in moving over the interval .03–.05 we are adding another two Type I errors in 100 replications. In this sense, the adjacent intervals shown in Table 2 may be regarded as equal intervals. For the most part, *S*s seemed to judge these intervals as approximately equal. One notable exception was the interval .05–.10 which .84 of *S*s judged to represent a greater decrease in believability than the subsequent interval .10–.15. The two-tailed *p* that this proportion exceeded the proportion judging the decrease in the first interval as either equal or less, was .004. None of the other analogously performed tests yielded a two-tailed *p* less than .36. Apparently, in going from the .05 to the .10 level there is a relatively more precipitous loss of confidence than is true at either higher or lower levels of significance. This is to be expected in that the .05 level has been traditionally considered by psychologists to be the upper limit of the region of rejection of null hypotheses.

D. DISCUSSION

The questionnaire asked for "expressions of degree of belief in research findings as a function of associated *p* levels." This request did not specify the exact criteria to use in defining degree of belief. The finding that sizeable proportions of *S*s had greater confidence in levels of significance based on an *n* of 100 suggests that most investigators not only use the probability of Type I errors as a criterion but consider (at least intuitively) Type II errors as well. If Type I errors only were of concern, the same degree of belief should be attached to the *p* for both *n*s. However, for a given *p* level

TABLE 2
PROPORTION OF *Ss* DIFFERENTIALLY JUDGING SUCCESSIVE INTERVALS OF LEVELS OF SIGNIFICANCE*

Decrease in confidence from:	Proportion of *Ss* judging as:			Than decrease in confidence from:
	greater	equal	less	
01–03	47	21	32	03–05
05–075	63	16	21	075–10
05–10	84	11	5	10–15
10–15	52	16	32	15–20
15–20	37	52	11	20–25
20–25	26	58	16	25–30
20–30	37	58	5	30–40
30–40	26	69	5	40–50

* For convenience, decimals have been omitted throughout the table.

the Type II error decreases with increasing n; thus the power of the test increases in that Power = $1 - p$ (Type II errors). Objectively speaking, one would expect greater belief attached to the larger sample size at every p level because for any p level, all samples of any size have the same probability of a Type I error but the probability of a Type II error is less for the larger ns. Thus, the statistical test using larger ns has a greater tendency to detect real differences if they are present.

In view of the emphasis, both in the literature and in elementary textbooks of statistics, on the .01 and .05 levels, we expected both of these points to show "cliff" characteristics, i.e., sudden drops in confidence in p levels just beyond these points. The .01 level did not show this hypothesized cliff effect but the .05 level did.

A major problem of the present study was the selection of psychologist *S*s from a single locus. The Ph.D.s held by the faculty member *S*s were awarded by state universities from the east and west coasts and the midwest. Points of professional origin, then, were well distributed. Most of the faculty and about half of the graduate students serving as *S*s were regular contributors to the psychological literature. While we cannot argue strongly for the representativeness of our sample, it appears to be at least a relevant sample. Nevertheless, our study should be replicated with another sample of psychologist *S*s.

E. SUMMARY

A total of 19 psychologists (graduate students and faculty) rated their degree of confidence in a variety of p levels for each of two assumed sample sizes. The relationship between degree of confidence and magnitude of p level appeared to be exponential regardless of sample size assumed and type of *S*s employed. *S*s had greater confidence in a given *p* level when it was associated with a larger sample size suggesting that investigators use the probability of both Type I and Type II errors as criteria of "belief." Graduate student *S*s tended to place more confidence in given *p* levels than did faculty members. For 84 per cent of the *S*s, the .05 level had cliff characteristics manifested by a relatively more precipitous loss of confidence in moving from the .05 to the .10 level than was true at either higher or lower levels of significance.

REFERENCES

1. GAITO, J. Statistical models versus empirical intuition: interactions in analysis of variance techniques. *J. Gen. Psychol.* (in press).
2. MCNEMAR, Q. Psychological Statistics. New York: Wiley, 1955.
3. RAY, W. S. An Introduction to Experimental Design. New York: Macmillan, 1960.
4. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
5. SNEDECOR, G. W. Statistical Methods. Ames: Iowa State College Press, 1956.

Department of Social Relations
Harvard University
Cambridge, Massachusetts

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 39-41.

EXPERIMENTAL RECOLLECTION OF DREAMS*[1]

Centro de Psiquiatria Experimental, Universidad de Chile

JULIO M. DITTBORN

It is a known fact that sleep can be interrupted by a perception of special significance among others of the same intensity. Experimental evidence of this phenomenon has been reported (1).

Starting from the above mentioned experiments I asked myself if it would be possible to interrupt sleep while dreaming by using an adequate stimulus related to this special situation.

I tried the experiment on myself by adapting a Grunding-Niki tape recorder with a closed loop which permitted a sentence to be heard continuously six to seven times per minute. Some nights before turning my room-light off I recorded the following sentence: "I am going to awake while dreaming" which I was to hear while sleeping throughout the night. Other nights I recorded a meaningless phrase of the same number of syllables which I was to hear under the same conditions. In a fifty-one-day period I submitted myself to nine experimental nights. On six of them the meaningful phrase was heard.

I started the experiments with the meaningful sentence. Every night I wrote down the time when the light was turned off, the moment of each awakening and the pertinent observations. The recording volume was adjusted so as to permit the initiation of sleep while hearing distinctly the corresponding sentence.

Contrary to my previous sleep habits before initiating these experiments, I had during the experimental period as well as afterwards frequent spontaneous awakenings during my nocturnal sleep.

During the experimental period I woke up all nights on which I heard the tape recording no matter which phrase was being played.

I woke up five times during the four nights on which the meaningless sentence was played. At none of those awakenings was a dream remembered.

I woke up ten times in the five nights in which I heard the phrase " . . . to

---

* Received in the Editorial Office on August 13, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported by the Faculty of Medicine, University of Chile (Grant 60-14) and by the Rockefeller Foundation (Grant 60038) under a joint program.

awake while dreaming." After four awakenings I could remember my dreams very distinctly; after five of them I remembered I was dreaming but I only could remember partially the dream text; after one awakening on these experimental nights I woke up without remembering anything at all.

The day following the experimental nights I felt unusually ill-tempered, nervous, and with a tendency to tremble, mostly in my hands. Owing to this fact I interrupted the tape-recording after one, or at the most two awakenings. I could also appreciate during this period a tendency to postpone the experiments as well as some reluctance to perform them.

As mentioned above, I also developed a tendency to wake up spontaneously on "free" nights. These spontaneous awakenings were more frequent during the free nights immediately after an experimental night. Five months have elapsed since the last experimental night occurred, and I still have a tendency to wake up spontaneously some nights. Moreover, I have developed a greater facility to remember my dreams.

A curious phenomenon happened after awakening during two of the nights in which I was listening to the meaningful phrase. The first time this happened, apparently when fully awake, with my night-table light on, I was hearing the sentence from the recorder in a completely distorted manner. At that moment I thought something had gone wrong with the tape. To my surprise, when I turned on the recorder the next morning there was the phrase neatly and exactly expressed as it had been registered by myself the night before.

Six days afterwards during a second experiment with the meaningful phrase the same phenomenon occurred. The distortion was this time quite different from that which occurred the previous time. I heard the changed text during seven minutes until I clearly understood the verb of the correct phrase and I could then rebuild the meaningful sentence.

The possibility was considered that the distortion could be the consequence of voltage fluctuations during the experimental nights when this phenomenon was observed. To obviate this, I submitted the tape recorder to wide variations of voltage before two observers while playing the significant phrase. Both said they could understand vaguely the phrase text at both extremes of the voltage range studied; in no instance did they hear the sentence as I had distorted it on the two occasions reported.

It would appear interesting to refer the following partial text of one of the dreams clearly evoked during one of the experimental nights:

> I was riding in a cab in Rio de Janeiro, a city which I had to leave the next day. I was writing desperately in a notebook phrases which I heard

> from a place I could not locate. At a given moment the notebook was filled with my notes and I continued my writing on the walls of the cab.

The stimulus distortion on one hand and the experience of accomplishing in a dream the task I was supposed to do upon awakening on the other, seem to uphold Freud's view that mental life during sleep tries to protect the sleep function from possible awakenings.

The first awakening during the experimental nights occurred, on an average, 138.5 minutes after the light was turned off. When meaningful sentences were played this average was 119 minutes and 158 minutes when the phrase was meaningless.

REFERENCE

1. OSWALD, I., TAYLOR, A. M., & TREISMAN, M. Discriminative responses to stimulation during human sleep. *Brain.*, 1960, **83**, 440-453.

Casilla (P.O. Box) 6507
Santiago, Chile

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 43-47.

WORK ATTITUDES OF WARD ATTENDANTS FOR THE CRIMINALLY INSANE*[1]

Educational Testing Service

MIRIAM WENDT AND HERBERT GERJUOY

Most authorities agree that in the constricted social milieu of a back ward, the attendants are the mental patients' most influential daily personal contact (3, 4, 5, 6, 7). Therefore, it would appear to be desirable that attendants' work attitudes be positive. Since back ward attendants work with passive, autistic, treatment-resistant patients, who may appear to be considered hopeless by the hospital professional staff, it seems likely that their work experiences predispose the attendants to perceive such patients as hopelessly ill. Thus, work experiences on chronic wards may tend to make aides become less optimistic about the very sort of patients with whom they deal. Support for the hypothesis that an attendant's attitudes are affected by the type of ward on which he works came from Gerjuoy, *et al.* (2). They reported that attitudes of acute ward attendants improved with time, whereas attitudes of back ward attendants became less desirable as time passed. In their study, the attitudes of attendants on various types of wards were measured with a six-month inter-test interval, using the Attendant Attitude Questionnaire (AAQ) (1). There was no direct measurement of attendants' attitudes as a function of service on a particular type of ward. Moreover, although Gerjuoy, *et al.* were concerned with job adjustment, they did not consider such personnel variables as age, education, length of hospital service, and length of ward service.

A. PROBLEM

This study is concerned with the hypothesis that attendant' attitudes toward patients, mental illness and mental hospital work are dependent upon these variables: length of service in the hospital, length of service on a particular ward, age, education, and type of ward assignment.

---

* Received in the Editorial Office on August 16, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is based on a portion of an M.A. thesis submitted by the senior author to the graduate faculty of the University of Toledo. The authors wish to acknowledge the willing participation of the staff at Lima State Hospital. We are indebted to Drs. Harvard L. Armus and Irma R. Gerjuoy for their critical reading of the manuscript.

Particular consideration was given to the relationship between attendants' attitudes and the kind of ward on which they worked because such a relationship, if significant, would suggest the advisability of rotating attendants through different kinds of wards and, perhaps, the advisability of an attendant indoctrination program for the improvement of attitudes among back ward attendants.

B. METHOD

This study was conducted in 1960 at Lima State Hospital, a 1500 bed hospital for the criminally insane. The staff-patient ratio was less favorable than at other state hospitals in Ohio. A prison sentence often awaited the patient who recovered.

The AAQ was administered to all members of the nursing staff on duty during the two day administration period. A higher AAQ score means greater agreement of the participant's responses with the responses of presumably "ideal" attendants. Only ward attendants' questionnaires were considered in this study. Participation in the testing program was required by the hospital administration. Of 320 ward attendants on the hospital personnel role, 228 reported for testing.[2]

Before each questionnaire administration, the participants were told that the purpose of the questionnaire was to learn more about nursing staff job satisfaction and morale. Anonymity of participants was stressed. Attached to each AAQ was an information blank where the participants were to record the personnel data: age, sex, education, and length of service on past and present wards.

All wards at Lima State Hospital, except observation and medical-surgical wards, are named and classified according to security measures required, from greatest to least security measures: *strong, ambulatory,* and *working.* These were the ward types considered in this study.

C. RESULTS AND DISCUSSION

Sixty-eight questionnaires returned by attendants were eliminated: 25 because of incomplete information, and 43 because the attendants were not assigned to one of the three ward types considered in the study.

Intercorrelations among the variables were computed. No variable correlated significantly with AAQ score.

The correlation between length of ward service and AAQ score was separately computed within each ward type. The correlations were not significant. However, they were in the expected direction and their magnitudes were in the expected order: for strong wards, .149; for ambulatory wards, —.095; and for working wards, —.170.

---

[2] This proportion of attendant participation was consistent with the normal attendant rotation practices during the two day testing period at the hospital.

A multiple correlation was obtained with AAQ scores as the dependent variable and the other variables as predictor variables. This was .306, significant at the .01 level. However, among the predictor variables, only length of hospital service and age contributed significantly to regression at the .05 level. The regression coefficients for ward type, length of ward service, and and education were not significant. The multiple correlation, using just the two significant predicator variables, was .236, significant at the .01 level. It would seem that the young attendant with much hospital experience is the attendant with the best work attitudes.

The hypothesis that attendants' attitudes are affected by and are dependent upon their length of service in a particular type of ward was not supported in this study, except for the doubtful implications of the direction and the order of the values of the correlations between length of ward service and AAQ score within ward types. Gerjuoy, *et al.* (2) reported a not much stronger, although statistically significant, effect of ward type upon attendants' attitudes. It should also be noted that the multiple correlation found in the present study between length of hospital service, age, and AAQ score, though statistically significant, was not large, accounting for less than 10 per cent of the variance in the AAQ scores. Moreover, the finding is of little practical value to hospital administrators because young attendants with much hospital experience are necessarily a rarity.

Lima State Hospital is a hospital for the criminally insane, and differs from the general psychiatric hospital considered by Gerjuoy, *et al.* (2) in several ways. A large proportion of the Lima patients are classified as sociopaths. Such patients are considered to have poor prognosis and to be unusually resistant to the influence of their social enviroment. Security measures are required and always enforced; a guard attends each ward. One would not expect to find an aura of therapeutic optimism in such a hospital.

The grand mean AAQ score of the Lima State Hospital attendants was 400.7, significantly ($P < .001$) lower than the mean of 462 for the Toledo State Hospital attendants in the Gerjuoy, *et al.* study. The t for the test of the significance of the difference between the means was 14.86 with 190 *df.*[3]

---

[3] The last two AAQ items were not scored in the present study. The maximum possible score on these two items combined was 10. Therefore, a correction of 10 was made in computing the t when testing the significance of the difference between the means. The means themselves, as reported, do not reflect such a correction.

This large difference between means implies an appreciably lower level of morale in the Lima State Hospital sample. Perhaps, low morale so pervaded all wards that individual differences in ward climate could not emerge.

However, in the light of the present finding, and the bare significance of the Gerjuoy, *et al.* findings (1962), their recommendations for rotation of attendants through different types of wards should be taken cautiously, since extensive revision of hospital personnel practices should rest upon a firmer objective foundation.

D. SUMMARY

The hypothesis was investigated that attendants' work attitudes depend upon age, education, length of hospital service, length of ward service, and type of ward worked on. Particular consideration was given to dependence of attendants' attitudes upon length of service in a particular ward type, since Gerjuoy, *et al.* (2) reported such an effect. Ward types were defined by ward security measures and restrictions on patient mobility, namely: strong, ambulatory, and working. It was predicted that service on wards with less stringent security measures would improve attendants' work attitudes.

The Attendants' Attitude Questionnaire (AAQ), with an attached personnel information blank, was administered to ward attendants at the Lima (Ohio) State Hospital for the criminally insane. Within each ward type, AAQ scores were correlated with length of ward service. No correlation within ward type was significant, although the correlations were in the predicted directions, and their magnitudes were in the predicted order.

A significant multiple correlation was obtained, with ward type, age, education, length of hospital service, and length of ward service as the predictor variables, and AAQ scores as dependent variable. However, only age and length of hospital service were predictor variables with significant regression weights. Attendants' attitudes did not depend upon experience on a particular ward type. Considering the low magnitude of the significant multiple correlation, the practical usefulness of the significant predictor variables must be questioned.

Over-all, attendant morale was much lower at Lima State Hospital than it was found to be by Gerjuoy, *et al.* (2) at a general psychiatric hospital.

REFERENCES

1. BOND, J. G., MCDEVITT, R., ROSENBERG, B. G., GERJUOY, H., & BALOGH, J. K. Milieu therapy with chronic regressed female schizophrenics. Final report to Ohio Division of Mental Hygiene. Project 55. Toledo: 1960.
2. GERJUOY, H., ROSENBERG, B. G., BOND, J. G., MCDEVITT, R., & BALOGH, J. K. Mental hospital ward attendants' work attitudes and their work experience in "front" and "back" wards. *J. Gen. Psychol.*, 1962, in press.

3. GREENBLATT, M., YORK, R., BROWN, E. L. (& HYDE, R. W.) From Custodial to Therapeutic Patient Care in Mental Hospitals. New Russell Sage Foundation, 1955.

4. SANDERS, R., FITZGERALD, B. J., HOBKIRK, J., SMITH, R. S., & WEINMAN, B. Socal rehabilitation of the Chronic Mental Patient. First interim report to N.I.M.H. Grant OM-126. Philadelphia: Philadelphia State Hospital, 1959.

5. SCHWARTZ, C. G., SCHWARTZ, M. S., & STANTON, A. H. A study of need-fulfillment on a mental hospital ward. *Psychiatry*, 1951, **14**, 223-242.

6. SCHWARTZ, M. S., SCHOCKLEY, H. L., & SCHWARTZ, C. G. The Nurse and the Mental Patient. New York: Russell Sage Foundation, 1956.

7. SCHWARTZ, M. S., & STANTON, A. H. A social psychological study of incontinence. *Psychiatry*, 1950, **13**, 399-416.

Department of Psychology
University of Toledo
Toledo 6, Ohio

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 49-61.

THE PRIMARY TRAIT ANATOMY OF THE MMPI* [1]

Department of Psychology, New York University

THOMAS N. JENKINS

A. PURPOSE

This study was a serious attempt to consummate a primary trait analysis of the MMPI scales. These scales are based upon clinical cases classified into clinical types "according to conventional psychiatric nomenclature." Hathaway and McKinley (5) refer to the personality characteristics of these types as "symptomatic complexes," thus implying that the scales are composite measures of a set of traits. The central purpose of our study is to determine the trait loadings of these complex measures. Such an investigation should furnish answers to several important questions, among which are the following: (*a*) Can the scales be differentiated from each other on the basis of patterns of trait loadings? (*b*) Can any of these scales be grouped into trait families? (*c*) Do the differential trait loadings facilitate additional insight concerning the nature of the "symptomatic complexes" measured by the MMPI? (*d*) Do the trait patterns suggest needed revisions of the scales?

B. PROCEDURE

1. *Subjects*

Our *S*s were undergraduate students in a first course of psychology at New York University. The sample comprised 104 women and 151 men, ranging in age from 17 to 22 years.

2. *Test Materials*

All of the tests were given during the first three weeks of the semester. The booklet form was used to administer the MMPI. The assessment of primary personality traits was obtained by administering a global personality inventory (9) of 720 items of the sentence type, preceded by the following 6-point, frequency scale: *Always, Almost Always, Often, Sometimes, Hardly Ever*, and *Never*.

---

* Received in the Editorial Office on August 20, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The publication of this study was facilitated by a grant from the Andreas Foundation, Minneapolis, Minn.

The development and form of this global inventory has been described in detail elsewhere (10). It furnishes measures of 128 primary personality traits. These scoring categories were based on a long-term factorial investigation of personality organization. The responses to items of large experimental inventories were factor analyzed for primary factors. Then scoring categories were developed for those primary factors which had been thoroughly validated. Three procedures were used to determine the factor loadings of the items: (*a*) graphical rotation to simple structure; (*b*) quartimax rotation to simple structure by the method of Neuhaus and Wrigley (15) with an IBM 650; and (*c*), a principal axis analysis by Univac at the Institute of Mathematics, New York University.

An important difference between this global battery of primary categories and the MMPI scales should be distinguished. The criteria for the global battery are factorial, and hence *internal*, while those for the MMPI are *external* diagnostic criteria. The global inventory, in other words, was based on *correlations between the responses* to items by the *S*s while the clinical scales were developed by *correlating item responses with opinions* about the *S*s.

C. PRESENTATION OF DATA

Pearson product-moment coefficients were computed for the relationships between each clinical scale and 128 primary scoring categories. This was facilitated through the use of IBM equipment. In the tables which follow, the tabular data are limited to coefficients which are extremely significant, that is, at a level of .001 or better. In interpreting data such as these, the size of the coefficient is also important. Statistical significance merely shows the probability of a true relationship. But in dealing with large samples, a true relationship may be so small as to be relatively inconsequential. For this reason, only correlations numerically greater than ±.31 are reported in the tables, that is, the coefficients for these traits which account for 10 per cent or more of the variance in the MMPI scales.

A routine presentation of the relationships between each clinical score and 128 trait scores would be rather baffling and frustrating for interpretative purposes. This confusion can be resolved in part by a rational system of classification, such as factor analysis, of the trait scores into a small number of categories. When the primary trait scores were factor analyzed, there emerged eight well-defined factors of the second order, hereafter referred to as superfactors. This analysis enabled us to classify the scores factorially into eight syndromes of primary traits. An extensive discussion of these second order components of human personality is published elsewhere (11). A major portion of the significant relationships between the clinical types and primary traits exists within the three superfactorial syndromes, which are most closely related to self-adjustment. The eight superfactors, which emerged as a result of the factor analysis of the trait scores, will now be described briefly.

1. *Adrenergy*. Usually characterized by a tendency to become unhappy. The term, adrenergy, was introduced by the writer (7) for theoretical reasons, since high scores seem to indicate sympathetic dominance, that is, a tendency to not only feel dejected, but also to feel sluggish, inferior, shy, tense, anxious, insecure, distractible, and habitually fearful. Table 2 contains a factorially pure syndrome of adrenergic traits, that is, the projections on all other factorial axes are small.

2. *Cholinergy*. Referred to as cholinergic because high scores suggest parasympathetic dominance, that is, a resilient tendency to be happy, and to feel energetic, enthusiastic, secure, and optimistic. A relatively pure syndrome of cholinergic traits is presented in Table 3.

3. *Sensory ascendance*. A superfactor closely associated with sensory phenomena. This syndrome comprises two related sets of traits, (*a*) the orexes or desires for sensory stimulation; and (*b*) a tendency to experience internal disturbances to relevant external stimuli (cf. Table 4).

4. *Aggressive egocentricism*. Socially maladjustive behavior often characteristic of a domineering, selfish, and quarrelsome individual (11, List 8).

5. *Self-sufficient intellectualism*. A dominant tendency to be theoretical, radical, unconventional, and independent (11, List 6).

6. *Realism*. High loadings characterize an individual who is able to present a realistic or objective account of himself (11, List 5).

7. *Dependability*. (*Socially approved goal orientation*). High loadings in this syndrome indicate a tendency to make acceptable adjustments to one's social environment, both vocational and non-vocational (11, List 7).

8. *Social reticence*. A tendency to keep not only one's feelings but also one's personal problems to oneself (11, List 10).

In the tables which follow, mean superfactor loadings are presented for each trait. These average loadings (correlations between the trait scores and the superfactor) are based on four diverse samples (323 undergraduate men, 200 basic airmen, 351 undergraduate women, and 324 retail saleswomen).

D. DISCUSSION OF RESULTS

1. *The Validity Scales*

Meehl and Hathaway (14) have suggested that the L, F, and K scales, although measuring no specific trait or clinical type, nevertheless belong to a family of variables. Some of the literature about the MMPI, however, shows that these distortion scores are related to various traits. Retholtz (16), for instance, found that the F score was related to certain physiological indices, such as Schneider's index, Darrow's automatic index, systolic blood pressure, and heart rate.

Table 1 reveals that all three of these scales are highly related to a large number of traits in the adrenergic syndrome. High trait loadings exist in the other syndromes, but about 70 per cent of such loadings are concentrated in the adrenergic syndrome. Since the greatest concentration occurs for K, this suppressor scale could be utilized to obtain a crude measure of adrenergy. These scales are related also to diverse sensory phenomena, as shown in Table 4, thus indicating that sensory ascendance is an important component of this family of distortion variates.

TABLE 1

CORRELATIONS BETWEEN VALIDITY SCALES AND TRAIT SCORES WITHIN THE ADRENERGIC SYNDROME

Column A Contains the Average Superfactor Loadings of the Traits

Test no.	Trait designation	A	L score		F score		K score	
			Male	Female	Male	Female	Male	Female
35	Indecisiveness	.65	—.32	—	—	—	—.40	—.48
37	Ruminative autisms	.70	—	—.43	—	—	—.50	—.64
41	Depression	.78	—	—.40	.50	.43	—.67	—.80
42	General anxiety	.74	—.36	—.40	.48	.40	—.69	—.50
43	Social insecurity	.50	—.32	—	—	.40	—.51	—.56
46	Hypersensitiveness	.70	—	—.43	.33	—	—.45	—.74
47	Self-commiserative autisms	.70	—.40	—.44	.50	.65	—.77	—.63
48	Cycloidism	.62	—.44	—	—	—	—.62	—.65
50	Lethargy	.55	—.35	—	—	.40	—.55	—
51	Nervousness	.73	—	—	.61	.47	—.46	—.61
53	Unwilling cooperativeness	.64	—.42	—	—	—	—.60	—.67
56	Need for dependence	.54	—.33	—	—	—	—.57	—.40
57	Self-algolagnia	.57	—	—.48	.41	.49	—.50	—.79
60	General (free floating) inferiority	.69	—.46	—.45	.49	.61	—.62	—.67
66	Face-to-face shyness	.54	—	—	—	.40	—.32	—.90
67	Grandiose autisms	.51	—.39	—.40	—	—	—.55	—.46
69	Seriousness	.45	—	—	—	—	—.37	—.56
84	Vocational insecurity	.69	—	—	.50	.70	—.61	—.71
93	Introversial distractibility	.68	—.38	—	.48	—	—.51	—.64
97	Spatial disorientation	.43	—.41	—	.33	—	—	—
99	Feeling of degradation	.60	—.32	—	—	—	—.40	—.40
100	Situational anxiety	.74	—	—.41	.40	.70	—.37	—.69
106	Intellectual inferiority	.52	—	—	—	—	—	—
107	Public shyness	.44	—	—	.40	—	—	—
112	Unpleasant reminiscence	.66	—	—	.40	.69	—.55	—.83
118	Temporal disorientation	.60	—.45	—	.47	—	—.60	—.40

These data disclose how the validity scales are related to each other. For instance, adrenergic and sensory ascendant factors seem to be *common* components for this family of variables. These results also lend support for the development of a *general* validity scale for the MMPI. An attempt in this direction was made by Gough (6) in his formulation of an "F minus K dissimulation index." The large number of high correlations in the adrenergic syndrome suggests that these scales are highly related to the superfactorial component of these traits. The K scale, for instance, is negatively loaded with practically all of the trait factors in the adrenergic syndrome. The total relational variance (sum of r^2) is much too great to account for the variance in the K scale without assuming a dominant role on the part of the common superfactor in producing these high relationships.

The importance of these findings for validative purposes will be discussed elsewhere. In this connection, it is of interest that the Jenkins global battery contains five measures of test attitudes which differ in nature from those of the MMPI. Some of these measures have proved to be important as predictors of individual retest reliability (12).

2. *The Clinical Scales*

The clinical scales resemble the validity scales in one important respect. They all contain a fairly large number of trait loadings in the adrenergic syndrome (Table 2). To an experienced clinician, the tendency for significant relationships to appear in the adrenergic area would not be surprising. The clinical scales, however, differ considerably from each other with respect to the other superfactors.

a. Depression (*D*) *Scale.* One of the interesting characteristics of this scale is its apparent bipolarity evinced by the fact that the adrenergic correlations are positive while the cholinergic relationships are negative. For instance, depression (tendency to be unhappy) correlates positively but buoyancy (tendency to be happy) correlates negatively with D. Previous research (7, 11) has shown quite conclusively that the adrenergic and cholinergic syndromes are statistically independent, that is, uncorrelated. Depression and buoyancy, therefore, are not positive and negative aspects of a bipolar trait but actually two independent traits. It follows that the D scale is not a bipolar variate, but a *mixed* scale comprised of two independent sets of traits. It is suggested, therefore, that in order to be maximally efficient, the D scale should be broken up into two subscales. Whether this could be done successfully would depend on the factorial composition of the MMPI items.

TABLE 2

Correlations between MMPI Scales and Trait Scores within the Adrenergic Syndrome

Column A Contains the Mean Superfactor Loadings of the Primary Traits

Test no.	Trait designation	A	Hs score		D score		Pa score		Pd score		Pt score		Sc score		Si score	
			Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
35	Indecisiveness	.65	.39	.65	—	.41	—	.50	—	—	.40	.50	—	—	—	—
37	Ruminative autisms	.70	—	.66	—	—	—	.58	—	.41	.75	.58	.52	.46	.33	.50
41	Depression	.78	.54	.71	.55	.50	.39	.46	.56	.48	.76	.79	.55	.54	.53	.60
42	General anxiety	.74	.46	.73	—	—	—	.42	.32	.50	.70	.60	.63	.50	.50	.63
43	Social insecurity	.50	.32	—	—	—	—	.46	—	.40	.50	.47	.42	—	—	—
46	Hypersensitiveness	.70	.37	.60	.48	.43	—	.40	—	—	.67	.62	.42	.40	.46	.55
47	Self-commiserative autisms	.70	.40	.70	.40	.35	—	.53	.32	.60	.75	.70	.52	.60	.47	.63
48	Cycloidism	.62	.56	.50	—	—	—	.41	.40	—	.77	.69	.55	.34	.42	—
50	Lethargy	.55	.40	.77	—	.44	—	—	—	—	.57	.41	.41	.34	—	.40
51	Nervousness	.73	—	.80	.39	.63	—	.48	.33	.54	.61	.69	.50	.60	—	.56
53	Unwilling cooperativeness	.64	—	.60	—	—	—	.43	—	.48	.48	.70	.34	.47	—	.41
56	Need for dependence	.54	.38	.46	—	—	—	—	—	—	.50	—	—	—	—	—
57	Self-algolagnia	.57	.36	.45	.35	—	—	.50	—	.44	.65	.54	.63	—	.45	.60
60	General (free floating) inferiority	.69	.60	.74	.43	.57	—	.70	.37	.75	.72	.70	.60	.65	.54	.67
66	Face-to-face shyness	.54	.36	.60	.52	.63	.32	.64	—	—	.66	.45	.47	.46	.62	.60
67	Grandiose autisms	.51	.40	—	—	—	—	.43	—	—	.44	—	.38	.49	—	—
69	Seriousness	.45	—	.60	—	—	—	—	—	—	.48	.60	—	.43	—	—
84	Vocational insecurity	.69	.37	.66	.37	—	—	.64	—	—	.51	.60	—	.55	.50	.68
93	Introversial distractibility	.68	.49	.56	—	.40	—	.45	.40	—	.59	.50	.61	.51	—	—
97	Spatial disorientation	.43	—	.37	—	.49	—	—	—	—	.32	.42	.35	.39	—	.57
99	Feeling of degradation	.60	.40	.57	.32	—	—	—	—	—	.50	.48	.43	.47	—	—
100	Situational anxiety	.74	.40	.70	.40	.32	.45	.43	.34	.62	.60	.77	.32	.73	.34	.60
106	Intellectual inferiority	.52	.33	.42	—	—	—	—	—	—	.50	—	—	—	—	—
107	Public shyness	.44	.35	.44	—	—	—	.59	.45	—	.60	—	—	—	.76	.52
112	Unpleasant reminiscence	.66	—	.60	—	.42	—	.51	.47	.53	.76	.71	.40	.67	.47	.53
118	Temporal disorientation	.60	.36	.56	—	.45	—	—	—	.42	.52	.53	.51	.44	.33	.67

The unipolar nature of personality traits is fully presented elsewhere (7, 11). But the bipolarity of the D scale warrants a brief discussion here. Cattell (2), in classifying the 4500 trait terms of Allport and Odbert (1) on the basis of meaning, found that the groups occurred in pairs, that is, "polar opposites to form bipolar continua." He also observed that some terms denoted a zero amount of the bipolar trait. When bipolar sets of trait words were utilized to rate individuals, factor analysis of these resulted in bipolar factors (3). In his comprehensive summary of studies of personality organization, Cattell (4) has cogently disclosed the fact that the bipolar view is a tradition among psychologists.

In our opinion, based on a careful comparison (7, 10, 11) of the results of content analysis and quantitative analysis, this traditional viewpoint is unacceptable. Items, which were grouped as opposites on the basis of *content*, proved to be statistically independent when the *responses* to the items were subjected to cluster or factor analysis. For example, depression and buoyancy, which are semantically opposite, proved to be independent psychological functions. Hence, high scores could be recorded for both depression and buoyancy for a given person. This means that an individual with high scores for both depression and buoyancy is one who also tends to react strongly to both depressing and happy influences (7).

b. Hypochondriasis, Paranoia, Psychasthenia and Schizophrenia Scales. These scales are grouped together because of certain basic similarities between the trait patterns. The resemblance is due to the concentration of significant trait relationships in two superfactorial syndromes, namely, the adrenergic and the sensory ascendant (cf. Tables 2 and 4). In certain other respects, however, the results for these four clinical scales differ from each other. It is worthy of note that *depression* and the *anxieties* (general or situational) are common characteristics of these four clinical types. In the case of Paranoia, there are no significant loadings for *lethargy, spatial disorientation*, and *temporal disorientation*. It may be that Paranoia differs in this respect from paranoid schizophrenia. It should be noted also that large sex differences exist within the adrenergic syndrome for Paranoia (cf. Table 2).

c. Hypochondriasis and Hysteria. Hypochondriacal and hysteric patients both complain about physical ailments. As the authors (5) of the MMPI point out, these common complaints sometimes make it difficult for the clinician to differentiate hypochondriasis from hysteria. The correlational patterns for the Jenkins global battery, on the other hand, differ markedly between these two clinical scales. As previously indicated, a large number of high relationships exist for Hypochondriasis within the adrenergic and sensory

TABLE 3

CORRELATIONS BETWEEN CERTAIN MMPI SCALES AND TRAIT SCORES WITHIN THE CHOLINERGIC SYNDROME

Column C Contains the Average Superfactor Loadings of the Primary Traits

Test No.	Trait designation	C	D score		Sc score		Mf score		Si score	
			Male	Female	Male	Female	Mean	Female	Male	Female
2	Face-to-face-initiative	.65	—.44	—.42	—	—.33	—	—.44	—.38	—.70
11	Buoyancy	.50	—.46	—.40	—	—	—	—	—.40	—.48
15	Social assurance	.50	—	—	—	—	—	—	—.36	—.56
17	Persuasiveness	.54	—.37	—.40	—	—	—	—.40	—.41	—.55
22	Building up others	.54	—	—	—	—.32	—	—	—.40	—
23	Cooperativeness	.54	—	—.40	—.33	—.32	—	—	—.40	—.46
26	Service-mindedness	.53	—	—	—	—	—	—	—.48	—
33	Emotional spontaneity	.66	—.45	—.56	—.45	—.60	—	—.45	—.44	—.66
63	Arbitrative initiative	.57	—.43	—.44	—	—	—	—	—	—
74	Social initiative	.62	—.60	—.33	—.42	—.41	—	—.60	—.60	—.57
77	Gregariousness	.49	—.42	—	—.42	—.41	—	—.48	—.62	—.48
87	Dynamism	.60	—.47	—	—	—	—	—.62	—.45	—.40
130	Sense of humor	.43	—	—	—	—	—	—.40	—.45	—.42

TABLE 4

CORRELATIONS BETWEEN MMPI SCALES AND PRIMARY TRAIT SCORES WITHIN THE SENSORY ASCENDANT SYNDROME
Column G Contains the Average Superfactor Loadings of the Traits

Test No.	Trait designation	G	K score		Hs score		Pt score		Pa score		Sc score	
			Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
8	Vestibular empathy	.59	—	—.46	—	.36	—	.36	—	.38	.34	.50
62	Psychosomaticism	.36	—.50	—.37	.38	.48	.55	—	—	—	.37	—
68	Undesirable compulsivenes	.44	—.40	—.77	—	.58	.38	.40	—	.40	.36	.32
70	Homotropism	.32	—.22	—.50	.37	—	.45	—	—	—	.37	—
72	Negative algolagnia	.45	—.32	—.44	—	—	.36	—	—	—	—	—
75	Gluttony	.31	—	—.55	—	—	—	—	—	.40	—	.53
78	Sexual romanticism	.23	—.32	—.45	—	.48	.32	.42	—	.49	.37	.39
86	Mucodermal orexis	.53	—	—.32	—	—	—	—	—	—	.34	—
89	Kinesthetic empathy	.44	—	—.56	—	.41	—	.32	—	.42	—	—
92	Vistorexis	.42	—	—	—	.47	—	—	.39	.40	—	—
115	Color shock	.45	—.42	—.64	—	.60	.51	.32	—	.35	.44	—
116	Naso-buccal epicurism	.24	—	—.42	—	—	—	—	—	—	—	—
117	Distancia	.47	—.48	—.70	—	.40	.37	.32	—	—	.40	.40
119	Olfactorexis	.46	—	—	—	.44	—	—	.34	—	—	—
121	Auditory orexis for non-musical sounds	.52	—	—	—	.32	.35	—	.48	.38	.32	—
123	Sexorexis	.28	—	—.43	—	—	—	—	.48	—	—	.40
124	Varietal love interests	.23	—	—	.32	—	—	—	—	—	.43	—
126	Transexual identification	.32	—	—.60	—	.41	—	.50	—	—	—	.32
127	Masochism	.43	—	—	—	—	—	—	—	—	—	—
129	Primary suggestibility	.34	—	—.38	—	.45	.32	.45	—	—	.40	.39

ascendant syndromes (Tables 2 and 4). Highly significant relationships, on the contrary, fail to appear for Hysteria in these two syndromes.

d. Psychopathic Deviate Scale. The distribution of the relationships for this scale throughout the various syndromes differs definitely from the other clinical scales. In the adrenergic area, it is theoretically interesting to find that *depression* and other related traits are closely associated with the Pd score. The relationships for traits, such as *cathetic obedience,* in the dependability syndrome are uniformly negative, while those for the aggressive egocentric syndrome are all positive. In the light of our present knowledge about psychopathic behavior, these results are not surprising. These data also indicate that the Pd scale is measuring what it is supposed to measure.

The authors (5) of the MMPI point out the danger of mistakenly diagnosing hypomanics as psychopaths. The relational pattern of primary traits for the Hypomania scale, however, is very different from that of the Pd scale.

3. *The Non-Clinical Scales*

Among the non-clinical tests which have been developed for the MMPI, the Masculinity-feminity and Social I.E. scales have been included in the standard battery. It is worthy of note that most of the high relationships for Mf occur for college women. This is especially true for Cholinergy (Table 3), Self-sufficient Intellectualism, Dependability, and Social Reticence. The reason for this outstanding sex difference for Mf is not clear. It may be due in part to differences between this modern college group and the so-called, normal group, which was used to standardize the MMPI. It also may be related to the fact that Mf is a bisexual scale. The writer (8), for instance, in a study of several hundred vocational interest items, encountered serious sampling difficulties, which induced him to develop two unisexual interest scales, that is, separate tests for masculinity and feminity.

The pattern of trait relationships for the Social I. E. scale shows that factorially it is very complex. Inspection of Tables 2 and 3 also reveals the fact that there are a large number of significant correlations in both the adrenergic and cholinergic syndromes. These coefficients, however, are consistently opposite in sign, thus showing that Si has been developed as a bipolar complex, despite the fact that Adrenergy and Cholinergy are statistically independent (7, 11). The separation of the positive and negative aspects of the Si test into two subscales would give a more effective indication of social potentials.

E. CONCLUSIONS

1. *Validity Scales*

The trait patterns for the validity scales show that they are crude measures of Adrenergy, this belonging to one trait family. K is the most heavily loaded in this respect.

2. *Clinical Scales*

Analysis of the Depression scale revealed that it is mixed, comprising two independent sets of traits, and hence, for maximal efficiency should be broken down into two subscales. The adrenergic and sensory ascendant patterns for Hs, Pa, Pt, and Sc indicate that these scales clearly constitute a trait family although they also differ from each other in certain respects. An important finding was that trait patterns clearly differentiate Hysteria from Hypochondriasis.

3. *Non-clinical Scales*

The Si scale is an adrenergic-cholinergic composite, thus comprising trait syndromes which are statistically independent. The separation of the negative and positive aspects of this category into two subscales would result in a more valid and efficient measuring instrument.

4. *Superfactorial Dominance*

Attention should be directed to an important uniformity in all the data. Within a given superfactorial syndrome the signs of the correlation coefficients were the same, that is, they were either all negative or all positive. This result indicates that the common component is more closely related to these abnormalities than the specific trait factors. This uniformity in turn suggests that these disabling abnormalities are constitutional in nature. The heavy concentration of significant and uniform relationships in several of the adrenergic syndromes serves to emphasize the importance of sympathetic dominance (7).

DEFINITIONS OF SCORING CATEGORIES

Brief descriptions will be presented here of those scores for which the one or two word trait designations fail to accurately convey the basic meaning of the scale. Some of the scoring categories, moreover, are relatively new in psychology.

8. *Vestibular Empathy.* Tendency to experience equilibral sensations in response to moving objects or living beings which are merely seen or heard, instead of depending strictly on the position or movement of the observer's body.
11. *Buoyancy.* A resilient tendency to be happy. *Buoyancy* is an antonym of *Depression*, but psychologically the two traits are actually independent of each other statistically (7, 11).
33. *Emotional spontaneity.* Tendency to express oneself in an uninhibited, natural, and enthusiastic manner.

39. *Cathectic obedience.* Liking to obey, involving a tendency to seek situations in which obedience is a highly acceptable form of conduct.
57. *Self-algolagnia.* Tendency to enjoy hurting or punishing oneself.
68. *Undesirable compulsiveness.* Feeling a strong need to do something which is socially unacceptable.
70. *Homotropism.* Tendency to identify more or less completely with members of the opposite sex, but not necessarily in a homosexual manner.
73. *Negative algolagnia.* Tendency to enjoy being hurt by others, or to derive satisfaction from mild pain.
78. *Sexual Romanticism.* Tendency to idealize relationships between the sexes. May be very unrealistic.
86. *Mucodermal orexis.* Tendency to enjoy and to seek stimulation of the cutaneous and membranous surfaces of the body.
87. *Dynamism.* Tendency to feel lively or energetic, with a sense of vigorous physical well-being.
89. *Kinesthetic empathy.* Tendency to have kinesthetic experiences (muscular or movement sensations) in response to activities which are merely seen or heard.
92. *Vistorexis.* Tendency to enjoy and to seek scenery, distant views, vistas and so on.
97. *Spatial disorientation.* Tendency to feel uncertain about directions or positions in space.
112. *Unpleasant reminiscence.* Tendency to be preoccupied with unpleasant memories.
115. *Color shock.* Tendency to become upset by bright colors or outstanding grays (black or white).
116. *Naso-buccal epicurism.* Tendency to be preoccupied with or interested in chemoreceptive stimulation, especially blends of chemical stimulation of the tongue, palate, and nasal membrane.
117. *Distancia.* Tendency to experience feelings of distance with respect to surrounding objects or social stimuli.
118. *Temporal disorientation.* Tendency to have a poor sense of time. Exhibited by a false sense of time intervals, distorted temporal memories, and so on.
119. *Olfactorexis.* Tendency to get great satisfaction from many different types of odors and from situations in which they occur.
121. *Auditory orexis for non-musical sounds.* Tendency to enjoy non-musical disturbances and to seek situations in which they occur.
123. *Sexorexis.* Tendency to be sexually appetitive, that is, to have a strong desire for sexual stimulation.
124. *Varietal love interests.* Feeling a need to love, not merely one, but several different persons.
126. *Transexual identification.* Tendency to identify with the cultural norms of the opposite sex.
129. *Primary suggestibility.* Tendency to be immediately induced to act in a specific way in response to the behavior of others.

REFERENCES

1. ALLPORT, G. W., & ODBERT, H. S. Trait-names, a psycho-lexical study. *Psychol. Rev. Monogr.*, 1936, **47**, 171p.
2. CATTELL, R. B. The description of personality: 2. Basic traits resolved into clusters. *J. Abnorm. & Soc. Psychol.*, 1943, **38**, 476-507.
3. ———. The description of personality: 3. Principles and findings in a factor analysis. *Amer. J. Psychol.*, 1945, **58**, 69-90.
4. ———. The Description of Personality. Yonkers: World Book, 1946.
5. HATHAWAY, S. R. & MCKINLEY, J. C. Minnesota Multiphasic Personality Inventory. Manual. (2nd Ed.) New York: Psychol. Corp., 1951.

6. Gough, H. G. The F minus K dissimulation index of the MMPI. *J. Consult. Psychol.*, 1950, **14**, 408-13.
7. Jenkins, T. N. The neutral theory of personality: new evidence and a revision of the argument. *Trans. N.Y. Acad. Sci.*, 1955, **17**, 315-30.
8. ———. How Well do You Know Your Interests? Form B-22, Manual. New York: Executive Analysis Corporation, 1957, 22p.
9. ———. How Well do You Know Yourself? Global Form A-42. T. N. Jenkins, 21 Washington Place, N. Y. Univ., N. Y. 3, N. Y., 1962.
10. ———. Measurement of the primary factors of the total personality. *J. of Psychol.*, 1962, **54**, 417-442.
11. ———. The second order components of human personality. *J. Psychol. Studies*, in press.
12. ———. Test attitudes as estimates of individual retest reliability. *J. Psychol. Studies*, 1962, **13**, in press.
13. McKinley, J. C., Hathaway, S. R., & Meehl, P. E. The MMPI: VI. The K scale. *J. Consult. Psychol.*, 1948, **12**, 20-31.
14. Meehl, P. E., & Hathaway, S. R. The K factor as a suppressor variable in the MMPI. *J. Appl. Psychol.*, 1946, **30**, 525-64.
15. Neuhaus, J. O., & Wrigley, C. The quartimax method: An analytical approach to orthogonal simple structure. *Brit. J. Statist. Psychol.*, 1954, **7**, 81-91.
16. Retholtz, G. An investigation of the relationship between involuntary behavior and personality traits. Doctoral thesis, N.Y.U., 1950.

Department of Psychology
New York University
New York 3, New York

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 63-90.

THE VISUAL APPEARANCE OF STATIONARY OBJECTS: SOME THEORETICAL PSYCHOPHYSICAL IMPLICATIONS OF THE REAFFERENCE PRINCIPLE*

Research Division, Columbus Psychiatric Institute, Ohio State University

SAMUEL J. M. ENGLAND[1] AND BENJAMIN PASAMANICK

A. PURPOSE

As an approach to the theoretical construct to be offered, we would like to present, first, the evidence for the fact that the environment external to the visual cortical system will be perceived for only a short time after the image becomes stationary with respect to the retina. That is, there is a disappearance of the steadily fixated visual image. Second is the evidence for the suggestion that much of the seemingly disparate data for nervous functioning that makes up the literature dealing with phenomenology finds unification when understood to be the consequence of *reciprocal innervation.* Perception would then be understandable in terms of two functions: the afferent nerve action potentials and the associated reafferent nerve action potentials, so that when nervous action becomes completed, immediate perception will have become simple and diminished in extent.

In particular we are interested in indicating the phenomenal similarities between the *Ganzfeld* experiments (including color disappearances), stabilized image experiments, experiments with binocular "rivalry," fluctuations of accommodation and of the pupil, experiments demonstrating centrifugal nerve impulses as inhibitors, and the demonstrations of the phenomenal stability of visual objects despite eye movements (Kornmueller).

Further, we would like to emphasize in discussion that while there is

---

* Received in the Editorial Office on August 21, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] It is not possible to overemphasize the importance of the assistance and encouragement that Dr. Donald R. Meyer provided to the first listed author of this paper. Dr. Meyer encouraged the inauspicious beginnings of this topic, and for over 18 months consented to criticise the many versions of this present paper. His careful guidance prevented this author from becoming overwhelmed in the morass of literature connected with this topic. Dr. Meyer is not however to be charged with the defects that may be inherent in this author's grasp of the subject matter, for he has not been consulted in these final drafts. The criticisms of the editors of several learned journals have assisted in bringing the notions formulated here into greater relief. The subject matter was discussed with Dr. Wolfgang A. Zeman who made a number of helpful criticisms. The careful and painstaking examination of the manuscript by Dr. Athan Karras is gratefully acknowledged.

experienced a disappearance of the stationary part of a visual image, the stationary round of nervous action is actively involved in shaping the content of the remaining perceived, non-stationary parts. That is, every perception has a contingent or variable part, and a uniform or stationary part and each is immanent in the other; and either mode cannot be considered separately to the exclusion of the other. These two modes of nervous activity—the uniform part and the variable part—have the relationship of complementarity to each other. These two ingredients of every perception are to be explained as being the consequence of *reciprocal innervation.* The term *reciprocal innervation* is used in the sense explicit in the theory of error-actuated electronic mechanisms. There is also a close analogy to the notion of the "back-emf" inherent in the operation of electrical machinery. It is the essence of nervous mechanism that it *operates* on the discriminated externality as though engaged in the solution of a mathematical problem in successive stages by the method of *operators* inherent in differential equations.[2]

Also, the suggestion is advanced that in view of the massiveness of the undiscriminated uniform part of nervous action, a theory of objects finds application within the context of psychological physiology. We also suggest that the disappearance of the fixated visual image can be a fruitful central theme to investigations elucidating the dynamic principles of classical or Pavlovian conditioning and some pertinent experiments are described.

B. THE EVIDENCE

There is a growing body of experimentation dealing with the recent discovery that a visual image, when subjected to stationary fixation on the retina, will be described as disappearing (9, 31, 43). It is observed by Krauskopf and Riggs (24) that: "When a target is imaged in such a way that it remains fixed in position on the retina, it is clearly visible at first, but it disappears with prolonged viewing." Under typical conditions of fixation there is always some movement of the eye and as a consequence the visual image is retained; without optical nystagmus to counteract the sensory nulling inherent in the neurological process, there ensues the disappearance of stationary visual images. It is by now well known that we habitually perceive by the method of successive differences. However, it is probably worthwhile to mention, that up to now the scientific information that may be derived from experimenting with the disappearance of fixated visual images has been almost exclusively available to such persons who may by chance or design have hit upon this perceptual experience during contemplation. The philosopher Bergson suggests that the differential calculus was almost assuredly discovered by this means. It is possible with sufficient practice to confirm for oneself that the visual field will precipitously disappear; that furthermore, among the astonishing effects that can be observed, all prominently vertical entities in the field may disappear independent of the horizontal elements and vice versa, and that this is freely controllable. One might be forced to conclude from this experience that the *fluxions* (fluctuations) of X plus the *fluxions* of Y make up an area of visual field after the manner of the integral calculus.[3]

---

[2] The reader will note that questions will arise in the course of reading that will be found explained later on in the text. No remedy for this has suggested itself, every relevant detail cannot be presented all at once. Reading through first rapidly to get the gist of the matter will be rewarding to a second more thorough examination. There will also be questions that are not answered at all, some of which are likely fundamental to our topic. Some indication will have been given of the place where further explanation is available; sometimes explicitly, sometimes only implicitly available.

Recently Chambers, Levitt, Carreras and Lin (7) have reported obtaining electrophysiological evidence which reveals direct neural pathways from the highest point of projection in the cerebral cortex centrifugally back to the peripheral somasthetic system. They observe that electrical activity of individual second order nerve cells, concerned with touch or tactile sensations in body parts, can be directly excited or blocked by neural activity originating at that cortical area to which these peripheral cells project. They express the feeling that such a system might operate so as to: "(I) refine the attributes of sensation, (II) organize perception, (III) form the early stages of mnemonic processes, (IV) set the receptive tone of the modality, or (V) initiate phenomena in the absence of peripheral stimuli." Indeed, all of these consequences are affirmed by this present theoretical formulation.

Figure 1 has been adapted from Ditchburn and Ginsborg (9) to convey the nature of the relevant eye movements. The principal movement to which reference will usually be intended is the rapid, star-step-like displacements, or flicks, of about 5 minutes of arc occurring at relatively *aperiodic* intervals averaging about 1 second in duration. It will be noticed that vertical displacements of fixation occur as well as horizontal, if somewhat less often.

The long period drifts are of sufficient magnitude to effect image nulling but this effect is of such a nature that discussion of this can be postponed.

---

[3] Those readers who might think this suggestion merely fanciful will be interested to consult "Newton's version of the calculus" in Bell (1940) p. 135 ff. Also the denunciation of B. Cavalieri in Bell, p. 121-122 for his envisagement of operations connected with infinitesimals. Cavalieri's suggestions are of greater interest when an exhibition of the geometry of disappeared visual images is a part of the personal repertoire of the investigator, than when looked upon as an historical curiosity.

Even under typical conditions of fixation the eye is always in motion. There are large slow oscillations, slow drifts up to several minutes of arc, and the inherent tremor of the higher frequency components of optical nystagmus,

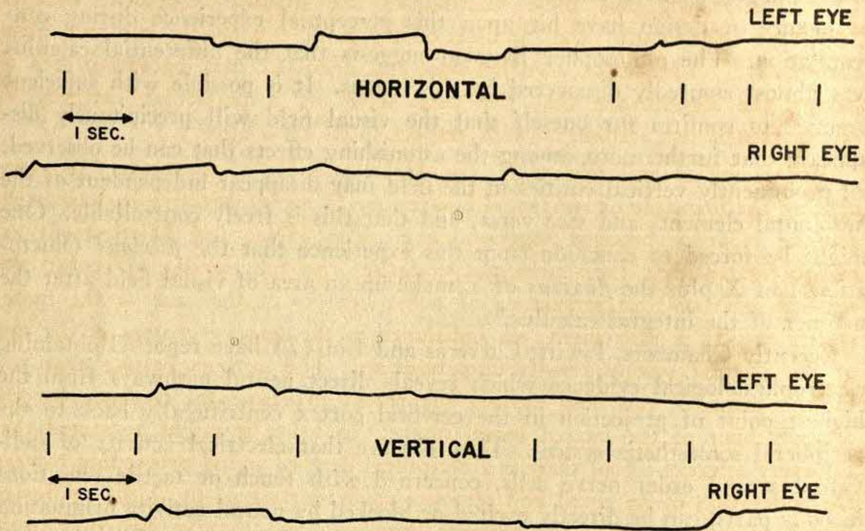


FIGURE 1

Freehand tracing of an electrograph record adapted from Ditchburn (1953). Lateral deviations of the upper pair of traces are proportional to the horizontal angular movements of the right and left eyes; the lower pair of traces show concomitant vertical movements.

i.e., rapid involuntary movements of the eyeballs from 10 to 20 seconds of arc. Ratliff and Riggs (30) report that component frequencies of eye movement range from below one cycle per second to around 100 cycles per second. The amplitudes vary from a few seconds of arc to several minutes of arc; all of these motions are relatively *aperiodic*.

Nystagmic movement of the eye is but one of a number of eye movements, which include pupillary fluctuations and accommodation changes. Stark and Campbell (36) report that the pupil constantly undergoes small fluctuations in aperture size. These oscillations are also of an *aperiodic* nature and not related to any inherent instability of the pupil system. They are rather small (± 10 per cent change in area of the aperture) which would seem to indicate that they do not prevent the disappearance of a stabilized image. Campbell and Robson (6) note that rapidly fluctuating patterns in the visual figures called to our attention by D. M. MacKay (25) are strongly related to accommodation. "The lens is constantly undergoing small (± 0.1 diopter), rapid (2–3 cps) fluctuations of power. This has been demonstrated objectively with an infrared high speed recording optometer. Arnulf *et al.,* (2) have shown that these fluctuations are mainly astigmatic, that is, "different segments of the circular ciliary muscle probably contract asynchronously." The terms *asynchronous* and *aperiodic* may be interpreted as having the significance that the time series of a particular category of observation is not stationary with respect to its own history.

The recent formulation of E. von Holst and H. Mittlestaedt (18) concerning the principle of *reciprocal innervation* is rather effective in accounting for the evidence brought together here. Von Holst and Mittlestaedt explicitly stated that their principle underlies the functioning of motor and sensory processes: an imbalance of nerve centers in the case of the motor system results in movement; an imbalance in the sensory centers in the case of the afferent system results in perception. The authors believe that von Holst's principle can help explain the recent experimental evidence for the disappearance of the fixated visual image. In brief, it might be stated that for every afferent signal there is a reciprocal derived signal proceeding toward the generation of an equal and opposite reafferent signal until the reafferent impulses exactly nullify the afferent impulses; and that if the equilibrium is modified by what might be termed a *constraining externality,*[4] a change takes place in the reafferent signal which at least partially annuls the constraint. It is our proposal that all central nervous system processes, both central and peripheral, may be shown to function in accordance with one unifying principle, i.e., the principle of reciprocal innervation. It is of central importance to note that Yves LeGrand (11) points out that in mammals the retina is an extension of the brain. And in another place, Polyak (26) states: "As is well known the retina is ontologically a part of the brain, and its structure resembles that of the central nervous system."

In a recent symposium on *sensory deprivation* (15) it has occurred to R. Held that the *reafferent principle* can account for many of the observed changes in perception characteristics of the "sensory deprivation" experiments. Although it is doubtful that this is a particularly good choice of terminology, the experiments conducted with the apparatus for effecting the disappearance of visual images might be termed "micro-sensory deprivation."[5]

---

[4] A constraint is a restriction on the geometry of motion. It reduces the number of degrees of freedom of the system. If the constraints depend on time implicitly, they are called *rheonomic*; if they do not contain the time implicitly, they are called *scleronomic. Rheonomic* may be thought of as flowing, time-like, implying relative motion. *Scleronomic* implies toughness, time-independence, therefore space-like, immobility.

1. *The Reafference Principle*

As von Holst (19) has put it:

> The whole of impulses which are produced by whatever stimuli in whatever receptors I shall call *afference,* and in contradistinction to this I shall call the whole of the motor impulses *efference.* Efference can only be present when ganglion cells are active; afference, on the contrary, can have two quite different sources; first, stimuli produced by muscular activity which I shall call *reafference;* second, stimuli produced by external factors, which I shall call *ex-afference.* Reafference is the necessary afferent reflection caused by every motor impulse; ex-afference is independent of motor impulses.

And again von Holst goes on to say:

> The initiated movement will continue until the reafference exactly nullifies the efference copy.

Concerning this point, with emphasis on motor responses, we may compare the words of Sherrington (34) p. 93:

> The correspondences support the view that the reflex inhibition (relaxation) and the reflex excitation (contraction) are part and parcel of one and the same reflex reaction; and that although opposite in direction they are coordinate reciprocal factors in one united response.

The scheme of afferent and reafferent nulling, suggested by von Holst (19), may be followed while referring to the signal flow diagram Figure 2. Consider any center, node, or terminal for example (z_1), which is supplied with

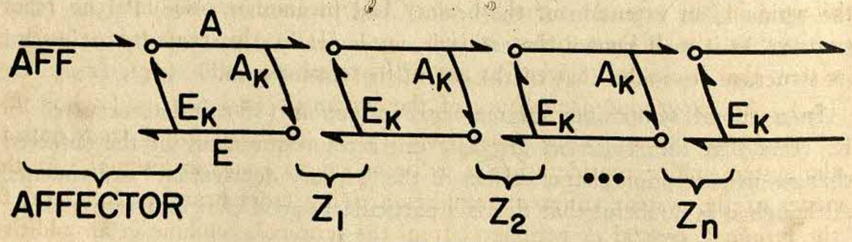


FIGURE 2

A signal flow graph representing the scheme of reciprocal innervation adapted from E. von Holst and H. Mittlestaedt (1950). There is an incident or right-going signal along the upper path, and a reflected or left-going signal along the lower path. In the steady-state condition the signals are balanced at the nodes or interfaces (Z_i).

---

[5] "Micro-sensory" in the sense of small, a small portion of the nervous system.

sensory signals from another center or node such as an affector (aff). The retina can be viewed as an affector; affectors may be any part of the nervous system consisting of many sensory inputs either of the same kind, or of more than one kind. The same principle can be said to operate whether viewed as a component or as an entire system.

There are several centers, terminals, or synaptic stations (z_2), . . . , (z_n) above the center (z_1). For a given sensory input (A) originating at (aff) there will result reafferent signal (E) activating the terminal (z_1). There will exist in addition an afferent *copy signal* (A_k) which spreads to the neighboring ganglion mass.[6] The reafferent signal (E) enters into reciprocal action with the signal (A). The afferent signal (A) and the reafferent signal (E) tend to neutralize each other. A signal originating at (aff) develops a residual at (z_1) and for that matter at other intermediate stations more central; this residual is conducted back extrinsically, sometimes outwardly to the most extrinsic center. On its way outward, considered branch by branch, the descending signal (E_k) will have been summed with the afferent signal (A_k). Only the difference (if any) subsequent to the initial volley will be perceived to ascend as a modified afferent beyond the lower order terminals. In the steady state, after all the *transients* have diminished, the systems in inward succession will be in a balanced or null condition. For example, if the affector (aff) should be disturbed there would develop a signal corresponding to this new disturbance. This signal, a perturbation from the mean nulled signal (A), will travel centripetally and modify the balanced, or steady state condition at (z_1), (z_2), . . . , (z_n). As the higher order centers detect the unbalanced condition peripherally, a corresponding signal will result in the (E) branches. A change is effected in the afferent signal at every intermediate station until no report arrives at the highest order station (z_n). It is the property of such a system that it can be "driven" at any of its terminals (aff), (z_1), (z_2), . . . , (z_n). An exogenous disturbance of the system at (z_1), . . . , (z_n) is just as surely an "input" as a signal from (aff). A signal input from outside the system under consideration is to be called an *exafferent* signal. At the nodes of the system, different impulses arriving from branches collateral to the terminal, central or peripheral from the terminal, combine in an additive positive or negative manner so as to increase each other or blot each other out. A strong signal ascending (A) may be nulled by its inverse, equally strong return signal (E); and a weak affector signal may be balanced at one or all of the stations with its weak reafferent nulling copy signal travelling back on branch (E). The foregoing considerations are discussed by E. von Holst and H. Mittlestaedt (18).

---

[6] "Copy signal" meaning reproduction or facsimile.

2. *Some Phenomena Examined in the Light of the Reafference Principle*

Historically one of the first experimental procedures which we can now explain as an instance of visual nulling is the *Ganzfeld* experiment. In this experiment the subject looks out into a diffused milky surface such as halved table tennis balls cupped over the eyes. This *Ganzfeld* situation offers the same signal to the eye no matter what direction is fixated, or how often the fixation or accommodation is changed. Cohen (8) reports his findings with the *Ganzfeld* phenomenon as follows:

> During the course of adaptation, 5 of the 16 observers reported a complete "black out" almost every time they were confronted with prolonged homogeneous stimulation. This was a unique experience which involved a complete disappearance of the sense of vision for short periods of time, and not simply the presence of a dark undifferentiated visual field.

The following description is representative: "Foggy whiteness, everything blacks out, returns again, goes. I feel blind. I'm not even seeing blackness. This differs from the black when the lights are out." A test of the homogeneity of the *Ganzfeld* is just this "feeling blind." The *Ganzfeld,* if it is truly homogeneous, is equivalent from the point of view of this paper, to a steadily fixated object of the Ditchburn, Riggs and Cornsweet kind. Anxiety and fear of blindness under somewhat similar conditions also were reported by Hochberg, Treibel and Seaman (17). Krauskopf (23) reports:

> . . . the well known fact that, although we do not ordinarily see the shadows of our retinal blood vessels, we can become aware of them if we cause them to vary by flooding the eye with diffuse flickering light or with light reaching the retina from various parts of the pupil.

Yet another empirical finding illustrating wider usage of this mode of thinking may be found in the work of Cohen (8) and Hochberg, Treibel and Seaman (17). Their investigation of chromatic neutralization has shown that a chromatic field will become chromatically neutral after prolonged inspection. They created a *Ganzfeld* condition of the visual field, by means of halved table tennis balls, wherein a homogeneous diffused illumination lacking any detectable pattern is regarded. The *Ganzfeld* condition is sufficient to effect image nulling under our principle, and indeed this occurred. Five *S*s reported: "Red surfaceless field followed by the total disappearance of the color within about 3 minutes." The report of four *S*s varied from a description of the field as dark gray, deep black, to a feeling of having gone blind. Hochberg reports that there were no significant recurrences of color during the 20-minute period except for slight and transient restoration of the sensation of color when the *S*s were instructed to attempt such a recovery by rapid eye movement. This result was contrasted with the situation reported by Helson and Judd (16) and testifies to the quality of Hochberg's *Ganzfeld.* Brief interruptions of the stimulus by blinking had no effect, but 2-second interruptions elicited the complimentary color, blue-green, for nine out of 10 *S*s. It is relevant to call our attention here to Shurcliff's report (35) of an irregular blotchy shape observed while viewing a uniform physical field.

The remarks of A. L. Yarbus (43) concerning the results of several experiments with fixated retinal images are also relevant:

> If the background was evenly illuminated, if its edges did not fall into the field of vision of the observer, if on the background itself there were no parts, which could move over the retina during movements of the eye . . . then after the cup blended into the background for the observer, the background itself ceased being perceived. The observer could say nothing either about the color or about the brightness of the background. Were some subject then to be brought sufficiently into the field of vision of the observer, perception of the background was re-established simultaneously, and with a light touch to the slide, that is, with a small shift of its image over the retina, it would be possible to instantaneously re-establish perception of the shutter.

It is possible without special apparatus to demonstrate to oneself the general character of the effects of chromatic neutralization by conducting Plateau's 1875 experiment as suggested to him by Wheatstone (33, p. 1063). A colored patch or rectangle can be fixedly regarded for 15 to 45 seconds in a well lighted room. Any bright color will do; bright red or yellow might be suggested. The effect can best be observed by regarding the rectangle with a silent and open mind, without seeking any particular result. When the observer first perceives this rectangle it will appear quite black or quite colored, with no particular detail within its boundaries. By fixating the object the observer will be able to note a change in hue within about 15 seconds. This can be checked by looking away briefly and then looking back. Also the internal region will be lighter than a border region that appears as a frame for the central region. The next stage will likely reveal a manifest halo of the complementary color around the border of the colored figure, by this time there will exist a well developed "after-image."

It is at this point that the demonstration can be said to begin. By carrying on with the fixation of the colored patch the observer may note that if he changes his gaze toward one corner or another of the figure, there will appear to be a milky film moving over the surface of the area, with exactly the same shape as the figure. There will be observed to come out of eclipse a portion of the patch from "under" the milky film which will once again appear quite colored. A 10 or more second continuance of the procedure of fixating the colored patch will serve to bring out one of the principle intended demonstrations: The milky whiteness of the "film" that now appears to cover the figure will become so compellingly obvious, with prolonged viewing, that one may encounter an urge to scrape at the figure with the finger nail to test the film's existence.

By this time the *after-image* of complementary color will be so firmly generated that one may look away to confirm the *after-image,* and gaze back to see the "milky film." It is suggested that the reader take this opportunity to demonstrate to himself the effects of *accommodation* on the apparent size of the retinal image. If one looks away from the page to a more distant point the apparent size of the *after-image* will enlarge very noticeably; the greater distance of the fixation of regard, the greater the apparent size of the *after-image.* We may now look back to the colored patch and fixate upon it for an interval once again. It is now appropriate to call attention to the detail that can be observed within the boundary of the patch. The nulling of the grosser aspects of the image have taken place; what is now compellingly transmitted centrally is the graininess in the texture of the figure. It is not suggested that this graininess of the texture could not be seen earlier and immediately upon glancing at the object, but rather that after the reafferent nulling has been effected the observer may be impressed with almost nothing else.

It is evident to the observer that all the grosser details are now fairly well nulled out. And the null-signal, the exafferent signal (the colored patch) and the reafferent signal (the complementary color) can all be observed. In particular the very fine detail that previously was only secondarily noticeable is now compellingly evident. It may be obvious that the characteristic color of this figure (the afferency) has been nulled by the reafferent signal, the complementary color. The milky film that the observer will be compelled to describe, is the null signal. The image that may now be seen is comprised of three signals. The observer may explain that while the milky figure is evident, and the colored figure is evident also, the complementary color is much less in evidence. Regard of a blank white area just below the figure will immediately reveal the complementary *after-image*. It is therefore clearly possible to witness the complementary image simultaneously with the colored and the milky white.[7]

Another instance taken from the domain of phenomenology, where our principle may serve to yield more than an interpretive description is the experience that has been called "retinal rivalry."[8] A simple figure such as a cross inscribed in a simple circle is viewed in an optical stereoscope. The cross is displayed to the left eye and the circle displayed to the right eye. At first the figure of a cross within a circle may be seen. With prolonged viewing the observer will notice the disappearance of first the circle, then the cross and these figures will continue to alternate in a more or less periodic fashion. This experience may be interpreted as being a method of demonstrating with very simple apparatus the disappearance of the fixated retinal image.

A number of motor-cognitive examples can be cited for their illustration as to how this generic principle would be expected to operate in the cognitive experience. Westheimer (38) reports his measuring of the accommodation of subjects in a dark empty visual field situation. The *S*s were asked to think of objects "quite close" to them or "far away." Appropriate .5 diopter accommodation changes resulted in the lens. The accommodations were transient, as a systems engineer might be led to anticipate. He disclosed a further result of possible theoretical import. "Two subjects were made angry by insulting remarks directed to them while the accommodation level was being tracked in complete darkness. In both cases a rise in accommodation level, lasting for several minutes, occurred; in one instance the rise exceeded 1 diopter." Westheimer comments, after investigating some of the cues for accommodation response to out-of-focus blurring of visual images,

> We find it difficult to believe that the accommodation response to visual stimuli does not involve the visual cortex in some way, much as the

---

[7] Some readers will find it provocative to examine the story of Okubo Schibutsu's bamboo painting in the light of the present psychophysical interpretation of this experimental result. An account of this classical incident can be found on p. 80 of Henry P. Bouie's essay on the laws of Japanese Painting.

[8] We feel that the image of a battle implicit in the word "rivalry" is wholely inappropriate; our thesis here is that both the *continuous* and the sensibly *discrete* or *particular* elements of a perception are different modes of one and the same nervous process. The error of conceiving the situation as one of "rivalry" is the error of the exclusive emphasis on the particular, discrete, and presentationally immediate parts of the perception: if a discriminated visual entity disappears and another is seen in its place, then the other must have aggressively displaced it. We are wont to emphasize that equally insidious errors await an exclusive emphasis of the continuous uniform infraparticular background of nervous activity.

> hand-eye task undoubtedly does. The high degree of versatility and rapidity of learning as well as the kind of detail in the optical image that has to be perceived make it unlikely that a simple path such as via the colliculus is involved.

Investigations of eye movement by Kornmueller (22) on himself described by von Holst give us evidence for the workings of reciprocal innervation, the reafference principle. Kornmueller anesthetized his eye muscles so that no intended eye movement could result in actual eye movement. Kornmueller found that for every intended eye movement to the left, which in this case was not a fulfilled eye movement due to the anesthetized muscles, there resulted the perception (Figures 3a) "the surroundings have jumped to the right." This can be explained, by means of the concept of the reafference principle, as resulting from a motor impulse stream (E) not nullified by

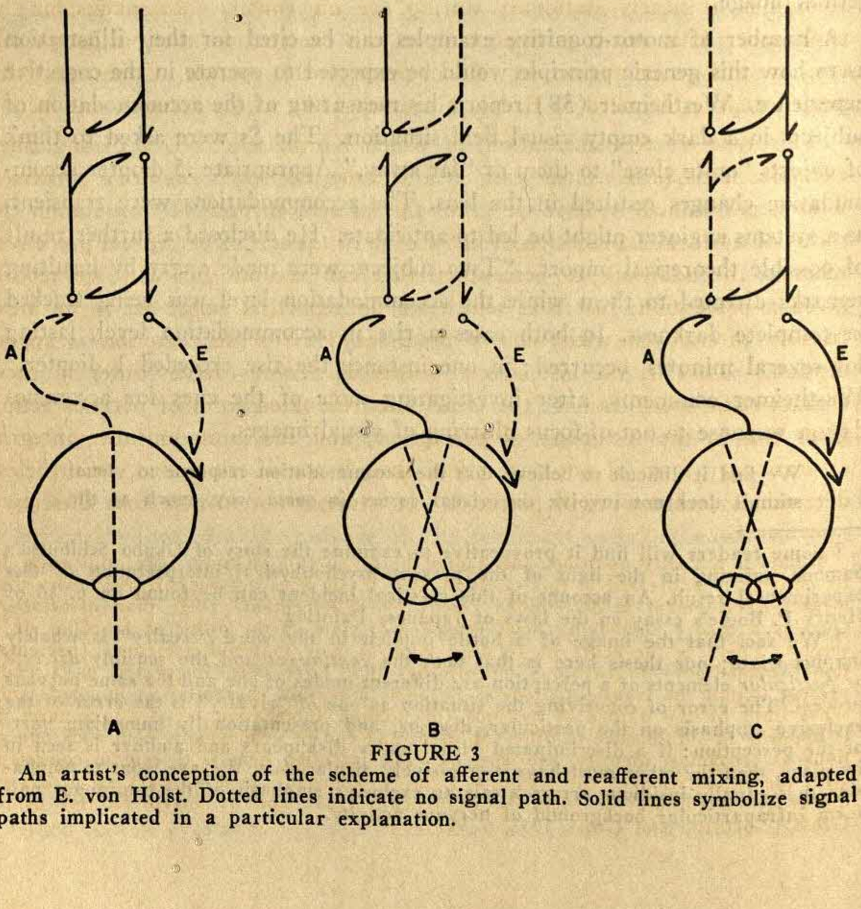


FIGURE 3

An artist's conception of the scheme of afferent and reafferent mixing, adapted from E. von Holst. Dotted lines indicate no signal path. Solid lines symbolize signal paths implicated in a particular explanation.

reafferent copy (A_k) proceeding back to a higher center (z_n) resulting in the above-mentioned perception: "The surroundings have jumped to the right." Next we turn the paralyzed eye mechanically to the right (Figure 3b). Both the motor intention (E) and efferent copy (E_k) are lacking. As the image moves across the retina, afference (A) is transmitted without nulling by (E_k) to higher centers. This results in the impression: "The surroundings move to the left." Now if we combine these two false perceptions, the perception of the returning unfulfilled "intent" resulting in a jump of the surroundings to the right, and image motion on the retina resulting in the impression of a jump of the surroundings to the left, we have the result that no movement of the surroundings is perceived (Figure 3c). The "correct" perception results from the cancellation of two "false" perceptions. Thus the psychological phenomenon of the perception of the surroundings as nearly stationary during eye movements (Raumkonstanz) is explained. This result is important to the explanation that, although our eyes are always in motion we seldom perceive this motion. Concerning this point attention is called to B. Kh. Gurevich (14) who alludes to this fact.

This type of experimentation, of which Kornmueller's is a distinguished example, is as near as one can, at this time, approach *cognitive* processes. This is particularly true if we grant this present theoretical formulation the status of a neural explanation of knowledge of the self of perceptual processes; wherein the additions to perception, that can account for why we report, given a perception, just that which we do report, is suggested to be inextricably interwoven with sensory processes. In a sense cognition is apperception.[9] That is, the two special forms of nervous action differ in their characteristic specifications: the clearly discriminated form of nervous action gives rise to the testimony of "perception," and the more uniform, integral or stationary form of nervous action, though not discriminated, nevertheless determines "what we will perceive." It is important to note that these two

---

[9] This is a knotty question, and cannot be given full attention here; this essay in its entirety is addressed to this question. Our thesis does not support the theory that mental development is determined chiefly by apperception rather than association. Ours is a theory embodying brief stretches of time of the order of fractions of a second. There is the sense of relating what is new material presented (presentational immediacy) to previous uniformities that have been integrated as already present within the transient process. There is a certain characteristic imposed upon sense data by the nature of nervous functioning. Integration is no vague term, it is opposite in significance to differentiation, and can be synonymous with such operations as recognize, identify, subsume, etc. The intention here is to remain consistent with the explanation of the technical term "prehension" as it appears on p. 235 of the essay on Philosophic Method (Whitehead, 1933) and as interpreted for us by Gross (Northrop and Gross, 1953, p. 927).

"modes" of perception are different modes of functioning of one and the same physical nervous mechanism.

3. *Implications*

Recently R. S. MacKay (26) has described a method that illustrates by analogy the effects that are operating in image nulling. "In essence, the method compares the image at one instant with its negative copy at a slightly later instant and displays only the difference between the two." Only those parts that fail to coincide fail to cancel.[10] MacKay describes the essential details of the system:

> If a motion picture is to be studied in this fashion, one first makes a negative print from it, developed to a gamma of unity. The print and the original are then shifted a few frames relative to each other and projected together . . . stationary parts of the scene will thus register as a uniform gray.

An appeal to this mechanistic analogy (Figure 4) permits us to suggest a functional significance for the "After-Image." The after-image may be looked upon as the physiological consequence of the reafferent nulling signal. The fact of its being seen "after" the afferent signal is no longer present merely testifies to the different transient characteristics of reafferency to that of afferency. We typically are most aware of the after-image when we regard an open nonstructured field. Such a phenomenon as the "flight of colors" exemplifies the ***homogeneous*** functioning of nervous action, having been set into disequilibrium by an initial impulse. Every afferent input to the nervous system entrains a response in the centrifugal nerve fibers which carry back an ***apposite*** reflective process in which all the elements hold each other in a state of mutual balance as in an electric circuit. There is no perception that is not constantly being modified toward this ideal. Plateau (1834) seems to have first explicitly suggested an active opposed process as the basis of the after-image; e.g., variations in the after-image, such as the "flight of colors," are due to an oscillatory process in the retina on its way to a ***state of equilibrium***.

Implicit in what has been set out above there seem to be some further consequences which we would like to offer for consideration. It is not possible to include all of the peripheral arguments which are, of course, required in firmly establishing associations and connections that are here suggested. Any device may be said to operate at an equilibrium when small transient disturbances do not permanently change its condition of operation. One type of equilibrium, which might be termed unstable equilibrium, may exist or be brought about in a system only to have the slightest disturbance permanently change the condition of operation of the system. It is quite possible for a system to possess more than one stable condition. Systems may have zero, one, two, or any number of stable operating conditions or states. The number of stable operating states and the rapidity with which they can change from one state to another may be one of the more fruitful ways of distinguishing the various sensory modalities. Let us consider the equilibrium condition.

---

[10] This type of monitoring of images of things can be made use of for the development of a "generalized movement function." The total light output of the slipped images, when monitored by a simple photo-cell, will be strongly related to a scaler or absolute value of total movement over the interval. This electrical signal might then be analyzed for its spectrum of periodic component elements, as well as for other statistically significant generalities.

If a system is in mechanical equilibrium, thermal equilibrium, and chemical equilibrium it can be said to be in *thermodynamic equilibrium.* If there are any changes in the state of the system, the system may be said to undergo a *process.* If, in the course of the process, the departures from equilibrium are negligibly small, the process is called a *quasistatic process.* That is to say, the process is very nearly a succession of equilibrium states. Actual conditions of ordinary perception are non-quasistatic; there are significant departures from quasistatic equilibrium. This fact does not affect the usefulness or importance of the notion of a quasistatic process as an ideal case in psychological dynamics. Once the transient differences are decayed, the system will continue in the steady state until acted upon by some different *exafferency.* This steady state has time-independent values at places in the system despite its being submerged in a dissipative process. This is the case for the *perceptual object*: Any other time-systems—which may be objects or partial objects—with which it can exchange energy are the surroundings of that object. We need not be concerned over our inability to write the equations expressing the mathematical relationships; at present it isn't necessary. A general concept of the nature of the relations between the properties of systems may actually be better conveyed by a graphical representation of the data. The experimentation of Pritchard (29), briefly outlined elsewhere in this text, may be thought of as a form of "graphical representation of the data."

The quasistatic equilibrium condition, which is brought about by the reafferentiation of sensory input, may be thought of as analogous to the "free body" of mechanics since we concern ourselves with a definite portion of nervous action bounded by a closed surface even though it may not be fixed in shape or volume. Some elements of any perceived configuration are always "stable," the fluctuations in novelty of the more permanent elements are noticeably slower than the "unstable" elements. The long term excursions—or fixed history—out of which events appear, form the *ground*. *Ground* is that element or those elements in a perceived configuration that are more nearly a balance of the *apposite* signals during the time interval under consideration: the infraparticular elements *patient* of the flux of transient experience. By *patient* is meant the condition of stationarity, time independence or uniformity in systems of nervous action that subsists throughout and extends over the transient part of a situation, providing *coordinative significance* to the inner events, and forming a primary feature in terms of which that perception may be interpreted. A flywheel may be thought of as *coordinating* the engine; or, to give a geometrical significance to *operators* familiar to mathematics, the XYZ inertial reference frame coordinates the $x^1y^1z^1$ non-stationary frame. The topic of monocular and binocular perception of depth would take us into the realm of the visual appearance of rapidly moving objects. Let us, for the present, confine ourselves to the subject of the visual appearance of stationary and quasi-stationary objects. It is to be recalled in this connection that only the transient part of the situation will be perceived, even though a continuous stream of impulses connects the higher centers with the lower ones, and some branches along the path will have reached a condition of perfect balance. To reiterate, the afferent opposite, or reafferent signal travels along the nerve tract, arriving at the synaptic stations where it mixes with the incoming afferent signal. The difference signal (if any) leaves the synaptic station to travel to a more central station where the principle of reafferent mixing is repeated. This suggests the principle that for every afferency there is necessarily an equal and opposite reafferency, during the steady state condition.

The perceptual object is poised as if between two negations of itself. Not all elements of the nervous system are as clear as the visual cortical process in exemplifying the terminal unity of nervous action, and the equipoise or balance that characterizes the perceptual object; but in principle all of the systems inherently proceed toward some specific eventual attainment.[11] The an-nulled perception is poised as a *time-independent process*, as if between two negations of itself. This is a restatement of the central characteristic of the reafference principle: a continuous stream of impulses connects the higher centers with the lower ones, even in the event that some branches along the path reach a condition of perfect balance or equipoise. The neural mechanisms are constantly operating on the incoming visual signal such that the apposite, afferent and reafferent, signals differ from each other by a transient amount.

---

[11] The objection has been urged against the foregoing suggestions that they are not experimentally oriented, that use is made of unfamiliar terminology and that the general tone is not *scientific*. It is true that the reader will not find the usual fare offered by empiricism; and some of the conclusions do not meet the specifications of that narrow theory of science, *logical positivism*. However they are offered as assumptions to be judged rather for results and consequences than for their initial familiarity. Professor Agnew (1942, p. 30) reminds us that: "To anyone who may object to making assumptions, we may say that much of the progress in science is due to men who have the courage to make assumptions, the good sense to make reasonable assumptions, and the ability to draw correct conclusions from the assumptions." One difficulty inherent in this present formulation has been suggested to be that the authors assume a multidiscipline competence on the part of the reader; that too much

As we have noted previously, the steady state condition issues in a balance with the constraining afferency, and no report will arrive at the higher centers. For this condition of nullity we may use the term *satisfaction*. Satisfaction is intended to be more than a figure of speech; this term is to have among its meanings that assigned to it by mathematics: the terminal unity of operation, the steady state condition toward which the neurological operations are proceeding.[12] With the introduction of adaptation toward the attainment of *satisfaction* order emerges. The nervous system differs little from other natural phenomena in this principal mode of its functioning. Were there no cyclical character to the flux of the constraining afferent action there would be no ordering of anything about this action. "The notion of 'order' always refers to something; that is, there is no order in the vague": Whitehead (1929). Typically we experience an *intersection* of the afferency and reafferency; that is: "The attainment is only partial, and thus there is 'disorder,' but there is some attainment and thus there is some 'order.' " The condition of satisfaction tends to stretch outwardly as it were, from a secular state of equilibrium into regions of non-equilibrium, an interweaving of change and permanence. For a more detailed elucidation of the concept of *order* an exegesis of the four grounds of the concept of order that can be found in Whitehead (41, pp. 127-167) is necessary. This cannot be included here.

A perturbation of a region of nervous activity that has proceeded to *satisfaction* will only transiently disturb the condition of satisfaction. This will, of course, result in the necessity for the nervous system to perform additional *work* on the incoming exafferent signal until the condition of satisfaction is

---

is expected of the reader. This may perhaps be so, however the subject matter seems to demand this minimum variety of analytical methodologies, otherwise no approximation to completeness can be realized.

[12] Those readers who may experience the apprehension or satisfaction that this type of formulation leads to a doctrine of psychophysical idealism may wish to be reminded that, insofar as this is the case, the situation is saved from an atavistic monism by the relativity interpretation of time, of multiple time-systems. What is not at present apparent is the importance to psychological physiology of interlacing multiple time-systems; in proportion as there are multiplicities of intersecting time-systems there is diversity—but not necessarily confusion—of experience.

again attained, there are no further disparities between afferent and reafferent signals, and the nervous energy has taken the form not directly associated with motion. In this connection it may prove useful to make a distinction between *smooth work* and *rough work*; as it will be generally agreed that work is being performed in any event, whether or not the transient departures from *satisfaction* are significant or nil. It is important to our thesis here that non-transient or permanent uniform departures from a given satisfaction imply a new and different satisfaction. A given system may have any number of satisfaction-states which require only smooth work to be sustained. However, the passage from satisfaction—which we may call a *state-point*—to satisfaction, will involve *rough* work, unless that mode of passing itself becomes instated as another *smooth* system of nervous action.

Such a schematization has inherent in it the principle of *least action* arising from the fact that every external constraint imposed on the system evokes a reaction that at least partially annuls the constraint. Consider the configuration of any dynamic system at a given instant of time somewhere in mid-career, and consider this *state-point* as having "specification numbers" to identify it. As the system changes its configuration with passage of time there are a series of state-points each corresponding to a new situation of nervous energy. The history of the system may be said to be epitomized by the set of state-points making up a "configuration space." Hamilton's principle of least action shows that the system moves along a path (successive state-points) so that the time-integral of action over the set of state-points is a minimum. One may find specific experimental exemplification or support of this principle by consulting the illustrations in Pritchard (29). He notes that a profile of a stabilized retinal image of a human head, for example, will be observed to disappear by discrete amounts which are invariably specific features or groups of features such as the face or the top of the head, or a distinctive coiffure, etc. And when monogram letters "break up," the fragmentary disappearance proceeds by discrete elements which are recognizable letters and numbers that constitute a subset of the group. Words projected for the stationary regard, become sensibly uncompounded in a similar manner. The subject will report that the disappearance proceeds in discrete steps of familiar words that constitute a meaningful subset of the original group or manifold. In Pritchard's experiment the word BEER provides the *specification numbers* of the first dynamic situation of the system, the next state-point has *specification* PEER, the next PEEP, the next BEE, and so on.

We would like next to advance the possibility of yet another point of view from which to characterize what we feel is occurring in the sensory process. We noted that on proceeding to completion, when the null has been effected, or neurological operation satisfied, the attainment is complete. We experience in such an event the *union* of afferency and its inverse, such that every element of the afferent set of nervous impulses is in one-to-one correspondence with the elements of the reafferent set. That is to say, the signals of the reafferent nervous pathways tend, by successive approximations, to become identical with the input energy spectra of the afferent pathways. It is of considerable interest in this connection to find considerations in matrix algebra, where the matrix product [A] × [B] may be zero even though neither [A] nor [B] is zero. The difference between the consequences of matrix multiplication and multiplication in scaler algebra—where in the relation $ab = 0$, either a or b is required to be zero—is not trivial and leads to useful statements about the nullity of a product matrix in terms of the ranks or nullities of its components, known as *Sylvester's Law of Nullity.*[13]

C. The Perceptual Object

That our usual perceptual experience does not attain the ideal, and is an *intersection* of the two pure modes of nervous activity does not modify the essential usefulness of our theory, which is an idealized linear theory obtaining only for small excursions from the stationary. We will next turn our attention to some of the clustering of evidence of nervous functioning that derive support from conclusions that have emerged from the Pavlovian experimental paradigm. There is general agreement in the literature concerned with classical conditioning that a nervous arrangement tends to form a fixed pattern which is sustained more easily as it is repeated under the same circumstances. This dynamic stereotype or systematic pattern can be maintained with less and less expenditure of nervous energy. Such a fixed pattern of nervous action develops an *inertia* which resists any alteration; and leads to the establishment of predominant processes, which act as *boundary conditions* for new stimuli. The cyclical and harmonic character of evoking externalities are arrested and sustained approximately to form a primary feature in terms of which the psychological experience may be interpreted.

This condition of inertia or rigidity is sufficient to call to our attention the theory of objects explicated in part by A. N. Whitehead (1920). For the present there is only space to suggest a few of the functional significances of the *perceptual object.* Perceptual objects form the conditions or *directrices* under which is arranged our perception; the dimly perceived permanent but essential factors governing perception are to be found as background in conscious experience. "The ingression of an object into an event is the way the character of the event shapes itself in virtue of the being of the object. Namely, the event is what it is, because the object is what it is . . . . " (40, p. 144).

---

[13] See Guillemin, E. A. The Mathematics of Circuit Analysis. John Wiley & Sons, Inc. New York, 1949.

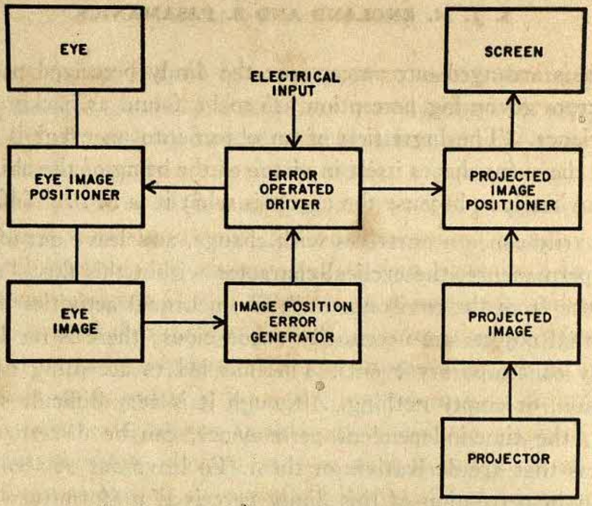
We ordinarily concern ourselves with change, and leave out of account the underlying permanence, the cyclical character within this flux, "mind's physical inheritance from the environment." The neuronal activities that form our pure perceptual objects are necessarily unconscious; there is no discrimination of absolutely contemporary events. This has led to accepting non-perception as annihilation, or empty nothing. Although it is very difficult some trace of the activity, the time-independent permanence, can be discerned as well as the sensations that are derivatives of them. To leave out of account the massive function in perception of this dimly perceived uniformity within perception is to leave out of account half of nervous functioning. When compared for massiveness, the enduring stabilities, the time-independent permanence in nervous action is enormous: what we perceive is subject to determinate conditions. We can no longer speak of perception as if it could exist in disjunction apart from a background. To suppose that there is a definite area of clear discrimination beyond which there is mere darkness, nothing of importance, may very well ignore the probability that the nervous system is complex enough to sustain systems of perceptions that presuppose inconsistent backgrounds. Perception is not fully described without reference to the time-independent uniformities, the *perceptual objects* that extend through the *specious present*.[14]

D. EXPERIMENTATION WITH THE PERCEPTUAL OBJECT

As to experimentation, the following program is offered as leading from the more speculative character of the middle toward a more familiar round of operations that some feel to be the sole and narrow realm of science. The degree and kind of movement of the image may be experimentally controlled by an apparatus hinted at in Figure 5. The power density spectrum of rhythmical eye movements may serve as a fruitful source of inference as to individual characteristics of perception, and changes in those characteristics. The visual display for either monocular or binocular regard may be set up in a manner suggested by Figure 6. The illuminated field is to be viewed

---

[14] The functional relationships that are signified to operate under the central notion of *specious present* may be found in James, W.; Dover 1950, p. 608 ff.



BLOCK DIAGRAM OF APPARATUS

FIGURE 5

Block diagram of apparatus. Such a configuration of parts as suggested by the system flow chart above would effect the fixation of an image on the retina of an observer, at the left, projected onto a screen at the upper right. Perturbations of the fixated image could be effected in a controlled manner at the electrical input to the system at the upper center.

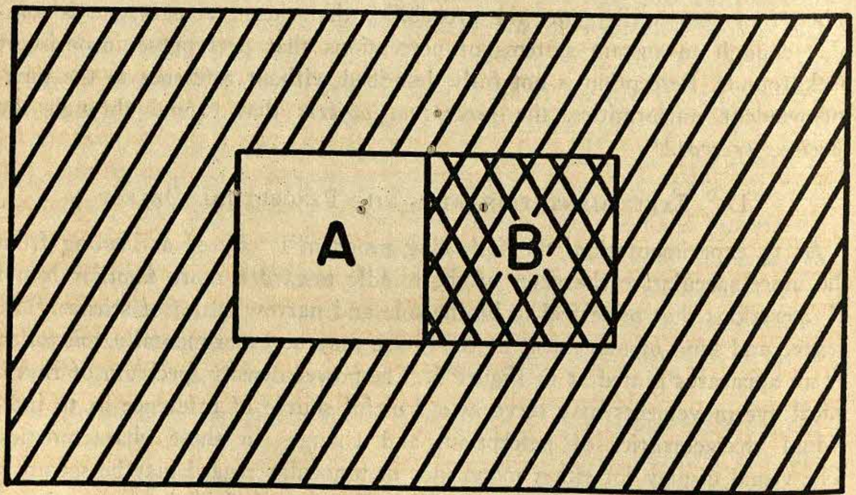


FIGURE 6

The patterns of areas represent images of differing energy level. A subject regarding these images fixated on the retina will note them to disappear. Systematic perturbations of this equilibrium condition can serve to reveal something of the functional relationships of energy transfer between neurological systems functioning in the uniform mode.

through an optical system perhaps operated by a Nystagmograph modified from that reported by Guillemin and Torok (13).

The necessary volume and field may be projected for regard by means of a multifield tachistoscope, a system of mirrors perhaps like that reported by Koletsky and Kolers (21). System A, Figure 6, is to be illuminated to a predetermined value of luminance; and the connecting system B is to have some other value of luminance, such as, for example, dark or zero luminance. The field surrounding the areas A and B may be 1/2 the luminance of A. As is now well established, the areas A, B and C will at first be seen, and then will, since fixated on the retina by the optical system, be noted to fade into quasi-static equilibrium. No report will be possible either as to the color or the luminance, shape or position of the objects now in the visual field. The luminance of both A and B may now be brought suddenly to one half the initial luminance of A. There will likely ensue a flow of energy from the field into the area B, and a flow of energy from the area A into the field. Eventually a new equilibrium state is reached where the entire system will again have become a uniform object. However, unless the energy flowing into the field from area A equals the energy flowing out of the field into area B, there will be a net flow into or out of the field and its energy will change. A change in the energy in the field will be perceived by the observer as a change in the apparent luminance of the field.

And then again, some one row of the configuration in Figure 7 might be projected with the same apparatus; the bodies will be noted to disappear. A small region or point source of light having a determined luminance can be caused to move across the nulled field in a precisely predetermined manner for the purpose of making note of any perceived deviations from its physical trajectory. A dimly illuminated graticule may prove necessary to the unequivocal determination of the trajectory. It is not possible to say at the moment but theoretical considerations would seem to predict that such a graticule, which would not be maintained stationary on the retina, and which would not therefore disappear, would itself be distorted in the regions near the disappeared perceptual object. In this manner, and with modifications from this paradigm, we would be enabled to study the relationships of *stationary objects* with *rapidly moving objects* or *events.*

An experiment can be performed that has bearing upon the theoretical principles underlying classical or Pavlovian *conditioning.* Figure 7 will serve to illustrate the account that we can make from our particular perspective of the conditioning paradigm. First, consider that the *manifold* of 12 squares is projected upon a screen for visual regard. The visual display for either monocular or binocular regard may be set up in a manner suggested by Figure 7. The image will be seen at first, and then it will be noted to disappear. No report will be possible as to color or luminance, shape or position of the objects now in the visual field. This sequence of events may be considered as the first trial or one replication. In subsequent presentation separated from the first by a period, say of 10 seconds, during which a blank screen or *Ganzfeld* is observed, the same sequence of events will be observed to occur: the image will be seen at first, followed by its disappearance. Should the time duration of the perceived image diminish, that is, at the first and most novel presentation, the duration of the perceived image will be the longest, and if at each subsequent presentation, the length of time that the image will be perceived is less and less, then we may say that *conditioning* or *learning* is involved. It may next be postulated that, should an audible tone or other sensory input be presented contiguous to or just prior to the presentation of the visual configuration, it is altogether likely that after a sufficient number of presentations, the temporal sequence of tone and visual display may be manipulated so that no perceptible visual image will occur whatsoever. Such an eventuality should be a sufficient criterion for the establishment of a con-

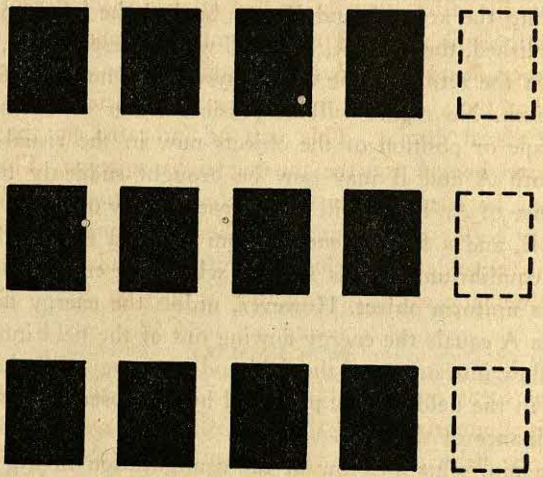


FIGURE 7

A manifold or society of colored patches. These definite areas, and additional ones, suggested by the dotted positions to the right, may serve as concrete evoking externalities for use in the study of classical conditioning utilizing the visual mode of sensory functioning for the determination of the dynamic relationships within the conditioning process.

ditioned reflex. The converse of this design may very well be important: can we condition the non-disappearance of a stationary object? With what physiological consequences?

Furthermore, the following line of reasoning suggests itself as more explicitly revealing the relationships that are involved: without a tone or any other cues except for a fixed time interval of, let us say, 10 seconds, the display of Figure 6 is to be repeatedly presented, such that the visual cortical nervous system is repeatedly called upon to reafferentiate the exafferent input to nullity or quasistatic equilibrium. After a nominal number of trials (to be determined by experiment) only four of the 12 squares would be presented; it is postulated that the visual cortical system will reafferentiate all 12 of the squares in the visual display. This will be indicated by the perception of the negative after-images for the missing eight squares. The four squares (any four) may be conceived as the *conditioned stimulus* related with the *unconditioned stimulus* of 12 squares. There are numerous other relationships and quantitative determinations that may be made using this elementary paradigm. An immense sea of investigation is open to us.

E. SUMMARY

Some of the recent experimental evidence for the disappearance of fixated visual images is summarized, and the suggestion is advanced that the formulation of a generic principle that can account for this result has been most effectively elaborated by E. von Holst and H. Mittlestaedt. The perceptual condition or state of nullity is suggested as a criterion of equality or *quasistatic equilibrium* near which experimentation can be performed with the goal of explaining how earlier phases of perception subsist submerged during the later ones, and how the cyclical character of evoking externalities are detected and sustained approximately.

In this physiological account of the visual cortical process it is presupposed that the totality of nature of interest to us contemporaneous with the conditions of external perception consists of two apposite streams of influence throughout the space of neurological events, balancing and intersecting and always influencing one another; that these apposite streams of nervous activity are paths of energy that cross and combine at centers or foci of dynamic balance.

The study of perception, as we are usually taught to understand it, concerns itself with discriminated neurological events and leaves out of account the dimly discerned uniformities: the neurological objects. This characterization is corrected by the suggestion that the complete experience is a transperceptual

experience wherein both of two pure modes of perception are immanent to the other. Each mode is fundamental to the characterization of the other. The exclusive predominance of either implies the non-existence of the other. This condition delivers us, with significant tempering of the concepts, to a complementarity relationship familiar to physics and elucidated by Rosenfeld (32). This is to say, that there is no total exclusion of the *smooth* modality characteristic of perceptual objects from the *rough* modality characteristic of perceptual events. The *perceptual object* may be thought to have the properties of a monochromatic wave in the ideal case, and the extreme opposite to the *perceptual event* having the properties of a particle. It is perhaps worthwhile to take cognizance of the fact that, as the concept is used in physics, *complementarity* denotes a relationship between concepts that are *logically* mutually exclusive; nevertheless both categories of process are necessarily to be found together in nervous action. As there is emphasis of either of the two extreme types of nervous action (pure perceptual events on the one hand, or pure perceptual objects on the other) we may note that the complementary one becomes less and less well defined, and is finally excluded. This may be thought of as an instance of the principle of procedure to simplicity by the diminution of extent. One may find specific experimental exemplification of this principle by consulting the illustrations in Pritchard (29).

There can exist at least two characterizations of what is going on: a characterization from the point of view of diversity, variation and separateness, and a characterization from the point of view or with the emphasis on uniformity, stationarity and unity. These two characterizations will often be at variance with one another, and mistakes can arise from the indiscriminate use of the contradictory concepts; but more properly they can be seen to mutually correct and assist one another particularly since they both apply to the same situation, denoted by the same time and place. Balanced between these two extremes may be found the usual perceptual experience. What category of balance and toward which extreme is a question for a theory of aesthetics.

Finally, experimentation is suggested whereby the mechanism of the *classical conditioning* experimental paradigm may be further elucidated. These seem implicitly pertinent to *learning* theory. It would appear at this time that the evidence suggests a *particle* or molecular theory of learning.

REFERENCES

1. AGNEW, R. P. Differential Equations. New York: McGraw-Hill, 1942.
2. ARNULF, A., DUPUY, O., & FLAMANT, F. Paris: *C. R. Acad. Sci.*, **232**, 349, 439, 1951.
3. BELL, E. T. The Development of Mathematics. London: McGraw-Hill, 1940.

4. Bouie, H. P. On the Laws of Japanese Paintings. London: Dover, N.Y.
5. Campbell, F. W. A high-speed infra-red recording optometer. *J. Physiol.*, **133**, p. 31P, 1956.
6. Campbell, F. W., & Robson, J. G. Moving visual images produced by regular stationary patterns. Nature, **131**, 362, 1958.
7. Chambers, W. W., Levitt, M., Carreras, M., & Lin, C. N. Central determination of sensory processes. *Science*, **132**, 3438, 1489, Nov. 1960.
8. Cohen, W. Spacial and textural characteristics of the Ganzfeld. *Amer. J. Psychol.*, **70**, 403-410, 1957.
9. Ditchburn, R. W., & Ginsborg, B. L. Vision with a stabilized retinal image. *Nature*, **170**, 36-38, 1952.
10. ———. Involuntary Eye Movements During Fixation. *J. Physiol.*, **119**, 1-17, 1953.
11. LeGrand, Y. Light, Colour and Vision. New York: Wiley, 1957.
12. Guillemin, E. A. The Mathematics of Circuit Analysis. New York: Wiley, 1949.
13. Guillemin, V., & Torok, N. Nystamograph for clinical use. *The Laryngoscope*, **68**, 120-130, 1958.
14. Gurevich, B. K. Conditioned fixation reflexes and certain aspects of the cybernetics of eye movements in man. *Biofizika,* **5**, 162-169, 1960.
15. Held, R. Exposure-history as a factor in maintaining stability of perception and coordination. *J. Nerv. Ment. Dis.*, **132**, 1, 26-32, 1961.
16. Helson, H., & Judd, D. B. A study of photic adaptation. *J. Exp. Psychol.*, **15**, 380-398, 1932.
17. Hochberg, J. B., Treibel, W., & Seaman, G. Color adaptation under conditions of homogeneous visual stimulation. (*Ganzfeld*). *J. Exp. Psychol.*, **41**, 153-159, 1951.
18. von Holst, E., & Mittlestaedt, H. Das Reafferenzprinzip (The reciprocal affects between the central nervous system and the periphery), *Die Naturwissenschaften*, **20**, 464-476, 1950.
19. von Holst, E. Relations between the central nervous system and the peripheral organs. *British J. Anim. Behavior*, **2**, 89-94, 1954.
20. James, Wm. The Principles of Psychology. **1**, Dover, 1950.
21. Koletsky, H. S., & Kolers, P. A. A multi-field electronic tachistoscope. *Amer. J. Psychol.*, **72**, 456-459, 1959.
22. Kornmueller, A. E. Cited in von Holst (1950) Eine experimentelle anasthesis der auseren augenmuskeln am menschen und ihre auswerkungen. *J. Psychol. u. Neurol.*, **41**, 354-366, 1931.
23. Krauskopf, J. Effects of retinal image notion on contrast thresholds for maintained vision. *J. Opt. & Soc. Amer.*, **47**, 740-744, 1957.
24. Krauskopf, J., & Riggs, L. A. Interocular transfer in the disappearance of stabilized images. *Amer. J. Psychol.*, **72**, 248-252, 1959.
25. MacKay, D. M. Moving visual images produced by regular stationary patterns. *Nature*, **180**, 849-850, 1957.
26. MacKay, R. S. Display of moving parts of a scene. *Science,* **130**, 244, 1959.
27. Northrop, F. S. C., & Gross, M. W. Alfred North Whitehead on Anthology. New York: Macmillan, 1953.
28. Polyak, S. The Retina. Chicago Univ. Press, 1941.
29. Pritchard, R. M. Stabilized images on the retina. *Scientific American*, **203**, 6, 72-78, 1961.
30. Ratliff, F., & Riggs, L. A. Involuntary motions of the eye during monocular fixation. *J. Exp. Psychol.*, **40**, 687-701, 1950.

31. RIGGS, L. A., RATLIFF, F., CORNSWEET, J. C., & CORNSWEET, T. N. The disappearance of steadily fixated visual test objects. *J. Opt. & Soc. Amer.*, **43**, 495-501, 1953.

32. ROSENFELD, L. Foundations of quantum theory and complementarity. *Nature*, **190**, 384-388, 1961.

33. SCHAEFER, E. A. Textbook of Physiology. **2**, Young J. Pentland, Publ. 1900.

34. SHERRINGTON, C. S. The Integrative Action of the Nervous System. Yale Univ. Press, 1906.

35. SHURCLIFF, W. A. New visual phenomenon: The greenish-yellow blotch. *J. Opt. & Soc. Amer.*, **49**, 11, 1041-1048, 1959.

36. STARK, L., & CAMPBELL, F. W. Pupil unrest: An example of noise in a biological servomechanism. *Nature*, **182**, 857-858, 1958.

37. STROUD, J. M. The Fine Structure of Psychological Time, *Information Theory in Psychology;* Proceedings of a conference on information flow, 1954. Glencoe, Ill.: Free Press, pp. 174-207.

38. WESTHEIMER, G. Accommodation measures in empty visual fields. *J. Opt. & Soc. Amer.*, **47**, 714, 1957.

39. WHITEHEAD, A. N. An Inquiry Concerning the Principles of Natural Knowledge. Cambridge Univ. Press, 1919.

40. ———. The Concept of Nature. Cambridge Univ. Press, 1920.

41. ———. Process and Reality. New York: Macmillan, 1929.

42. ———. Adventures of Ideas. New York: Macmillan, 1933.

43. YARBUS, A. L. New methods for investigating the functions of the various parts of the retina. *Biofizika*, **1**, 5, 163-165, 1957.

Columbus Psychiatric Institute & Hospital
University Health Center
473 W. 12th Avenue
Columbus 10, Ohio

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 91-99.

COMPARISON OF PHOSPHENE THRESHOLD TO VARIOUS VISUAL FUNCTIONS AND TO ELECTRICAL SENSITIVITY OF THE SKIN*[1]

The Training School at Vineland, New Jersey

JOHS. CLAUSEN AND RATHE KARRER

A. INTRODUCTION

From extensive phosphene studies in this laboratory it has become apparent that there are considerable individual differences in threshold magnitude. Under optimal stimulus conditions (20 cps frequency, .03 mL background illumination) thresholds vary from five to 40 μA in different individuals, but repeated testing over a long period of time, in a few cases up to five years, shows little variability of the threshold. While there is a fair amount of information with regard to the effects of stimulus conditions, such as wave form, frequency and light adaptation on the threshold, the question of the significance of the threshold magnitude in relation to other psychological or physiological functions is largely unanswered. It is this question that we have recently focused upon largely through correlational studies.

Previously we have investigated the relationship in mentally retarded *S*s between phosphene threshold and 92 variables, representing a variety of sensory, motor, perceptual and complex mental functions (5). Although a few significant correlations were found, they were small and in view of the large number of correlations computed, they were considered to have occurred by chance. A significant difference, however, was found between phosphene thresholds in retardates and normals. Evidence was presented against the higher retardate threshold being a secondary manifestation of low mental level. Also the data showed a strong indication of a relationship between phosphene threshold and per cent alpha in retardates with evidence of organic involvement.

The present study is an attempt to investigate in normal adult *S*s the relationship between phosphene threshold and other variables, particularly those which in various ways involve the visual system. The following measures were included in the study: phosphene thresholds at 20 and 50 cps,

---

* Received in the Editorial Office on August 25, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported by the National Science Foundation, Grant G-7029.

absolute threshold for light, brightness discrimination, CFF, perception of apparent motion (phi-phenomenon) and electrical sensitivity of the skin at 20 and 50 cps. This selection resulted from an analysis of the phenomenal experiences and possible physiological processes in the perception of the visual flicker which results from the electrical stimulation.

The absolute threshold for light may be assumed to reflect primarily the sensitivity of the visual receptors (photochemical processes) to photic stimulation, while the phosphene threshold may reflect the electrical excitability of the bipolar or ganglion cells of the retina. On this basis no correlation would be expected between the two thresholds. On the other hand, the light threshold may also be a measure of a neural excitability and the phosphene threshold may also reflect the sensitivity of photochemical processes. In both cases a significant correlation would be expected between the two thresholds but the result would be difficult to interpret by itself.

The phenomenon of phosphenes involves a discrimination of flickering light superimposed on the background luminosity of the screen which the subject is facing. If this element of light intensity discrimination is similar to traditional brightness discrimination the two measures should be correlated. Brightness discrimination, however, involves some sort of spatial quality which is not apparent in the perception of phosphenes.

Phosphenes at threshold intensities appear as a faint flimmer. This flimmer has a certain element of illusory movement which is most apparent at very low frequencies. It is possible that this movement-like component is an important cue at threshold intensities. It was therefore decided to include a measure of apparent movement. It was a difficult task to find a measure which isolated the movement-like element of the phosphenes. After having considered several alternatives, we finally decided that the phi-phenomenon was the most appropriate measure. If the movement-like component is an important factor for the phosphene threshold, a correlation would be expected between the phosphene threshold and the phi-phenomenon.

The subjective sensation resulting from intermittent photic stimulation and from alternating or pulsating current are so similar that it is possible that the same mechanism is involved in both. As a measure of sensitivity to photic flicker CFF is the traditional measure. There are, however, distinct differences between the two phenomena: critical flicker frequency is primarily concerned with temporal factors while the phosphene threshold is primarily a function of stimulus intensity. It should be further noted that fusion is not obtained for alternating electrical stimulation at frequencies which would definitely give fusion of photic stimulation. We have observed flickering phosphenes at frequencies as high as 120 cps. In view of these considerations we would not expect a relationship between phosphene threshold and CFF.

On the assumption that the phosphene threshold reflects the excitability of the nerve fibers in the retina, the question arises whether this sensitivity is specific for the visual system or if it reflects in general the neural excitability of the individual. The electrical sensitivity of a sensory modality other than vision should give evidence for such generality. Cutaneous sensitivity was selected because of the advantages with regard to electrode placement and general experimental procedure.

B. METHOD

1. *Procedures*

Phosphene threshold was recorded at 20 and 50 cps sine waves while *S* fixated a spot on a screen which reflected .03 mL. Five thresholds were determined for each frequency. Procedure in general, instrumentation, electrode placement and presentation of stimuli were the same as used in previous studies (4, 6).

Brightness discrimination was determined by binocular comparison of the two halves of a bisected circular target of opal glass, four cm in diameter in a darkened room. An incandescent lamp was the light source for each half, one of which served as the reference and the other as the variable brightness controlled by a General Radio Voltage Divider, model 1454 A. The Method of Limits was used with alternating ascending and descending series of equal steps of .0002 on the voltage divider. At each setting *S* was asked to judge which side was brighter. A series was completed when the *S* gave three consecutive judgments of brighter for the half which was initially dimmer. The PSE was determined from four series and transformed to milli-lamberts. The difference between this determination and the reference value constituted the score.

The absolute threshold for light was determined by using a 500 W slide projector as the light source, with a system of neutral density filters and a three log variable neutral density wedge. The light was focused and its intensity determined on an opal glass target with a Macbeth Illuminometer. A reduction screen containing a 1.5 mm aperture was inserted for threshold determination. Included in the arrangement was a photographic shutter which allowed a stimulus presentation of .5 seconds. By means of a dim red fixation point, the stimulus was made to fall upon the retina 45° from the fovea. A chin rest held *S*'s head in a fixed position 12 cm from the target. Since the phosphenes at threshold intensity are seen in the peripheral field, it was felt that absolute threshold for light should also be determined peripherally. *S* was dark adapted for 30 minutes prior to the test and the threshold was determined in an ascending series in steps of .04 log units. The total neutral density (N.D.) of all filters at threshold was the measure used since it is difficult to measure accurately light sources of this nature with the Macbeth Illuminometer. The mean of five threshold determinations constituted *S*'s score.

CFF was recorded by means of a Krasno-Ivy Flicker Photometer. Standard procedure with regard to viewing distance and room illumination for this photometer was adhered to. Starting at extreme frequencies, four ascending and four descending series were conducted in alternating directions. For each series *S* was asked to judge when the light changed from steady to flicker or from flicker to steady. The score was the mean of the eight determinations transformed to cps.

Phi-phenomenon was recorded by using two 1.5 volt lights (GE-112) mounted 6.5 inches apart through a 16 × 22 inch black cardboard background. The lights formed the base of an equilateral triangle, the apex of which was a white fixation point. The time sequence was achieved by a re-cycling arrangement of four Hunter Interval Timers so that the duration of the two light flashes was fixed at 0.1 seconds. The duration of the time interval between lights 1 and 2 and between 2 and 1 was varied equally in steps of .02 seconds. Four ascending and four descending series were conducted. *S* was seated seven feet from the display under 160 W fluorescent room illumination. He was instructed to judge at each duration whether the light flashes seemed to be discrete or whether there seemed to be a movement from one light to the other. The durations (in seconds) of the time interval for the transition points between moving and discrete flashes of four ascending and four descending series were averaged and constituted the score.

Cutaneous thresholds were recorded by using the same instrumentation and the same general procedure as the phosphene threshold. Two electrodes were mounted three inches apart on a plexiglass holder which was strapped to the volar surface of the right forearm three inches above the wrist. *S* was told to indicate for each stimulus presentation whether or not he felt a prickle, an itch or a sensation of warmth which were taken as threshold criteria. Ascending and descending series of stimuli of one second duration were presented in increasingly smaller steps as the threshold was approached. Five thresholds (mean of ascending and descending series recorded in μA) were determined for 20 and for 50 cps current.

To obtain an indication of reliability, all measures were determined a second time a few days later.

2. *Subjects*

A total of 14 adult normal *S*s, six females and eight males, were included in the study. Age ranged from 21 to 49 years with a mean of 31.6 years.

C. RESULTS

The means and standard deviations for the eight scores are presented in Table 1.

TABLE 1
GROUP MEANS AND STANDARD DEVIATIONS BASED ON THE MEAN OF THE SCORES FOR THE TWO SESSIONS
(N = 14)

		Mean	SD
Phosphenes	20 cps (in μA)	16.2	9.8
Phosphenes	50 cps (in μA)	116.6	44.0
Cutaneous	20 cps (in μA)	95.0	74.7
Cutaneous	50 cps (in μA)	161.0	108.0
Threshold for light	(in N.D.)	5.34	.20
CFF	(in cps)	35.75	2.01
Brightness discrim.	(in mL)	3.21	1.83
Phi-phenomenon	(in sec)	0.32	.14

The absolute magnitude of these means are of no particular concern here, since the primary problem is the relationship between phosphene threshold and the other measures. It is worth noting that CFF and threshold for light (which is of the order of $.1482 \times 10^{-4}$mL) show considerably less individual variability than the other measures and are the only visual measures which are directly correlated (—.53, significant at five per cent level). A comparison of cutaneous thresholds for 20 and 50 cps shows that the well established phosphene finding of a higher threshold at 50 cps than at 20 cps (3) has its parallel in cutaneous thresholds. The sensitivity of the stimulated structures in the eye, however, is appreciably greater than those of the skin.

Reliability of the measures were determined by correlating the results of the first and second session. These correlation coefficients are entered in Table 2. It is seen from the table that six of the measures are highly reliable. Cutaneous threshold at 20 cps barely reaches the five per cent level, but this low correlation is primarily due to three *S*s. By disregarding these *S*s the coefficient is increased to .92. The low reliability of the brightness discrimination cannot be traced either to particularly variable *S*s or to instrumental factors. It may be a result of the failure to standardize the dark adaptation of *S*s.

TABLE 2
RELIABILITY COEFFICIENTS BETWEEN TEST-RETEST MEASURES

	r
Phosph. 20 cps	.97
Phosph. 50 cps	.79
Cutaneous 20 cps	.54
Cutaneous 50 cps	.85
CFF	.83
Brightness discrim.	.31
Phi-phenomenon	.87
Light threshold	.88

1 per cent level = .78
5 per cent level = .53

The product-moment correlations between the phosphene thresholds and the other variables are listed in Table 3. None of these correlations reach significance. It appears, therefore, that phosphene threshold is not correlated to any of the other measures.

TABLE 3
PRODUCT-MOMENT CORRELATIONS BETWEEN PHOSPHENE THRESHOLDS AND OTHER VARIABLES

	Phosphene 20 cps	Phosphene 50 cps
Cutaneous 20 cps	.16	.00
Cutaneous 50 cps	.45	.35
CFF	—.35	—.17
Brightness discrim.	.01	.04
Phi-phenomenon	.32	.28
Threshold for light	—.06	—.16

5 per cent level = .53

Even though phosphene thresholds are not correlated with cutaneous thresholds there is a high correlation between the two frequencies for both measures (phosphenes 20 and 50 cps = .93 and cutaneous 20 and 50 cps = .83, both significant at the one per cent level).

D. DISCUSSION

The purpose of this study was to investigate whether or not phosphene threshold is related to various measures which involve the visual system or to electrical sensitivity of another sensory modality. Expectations with regard to these relationships were presented in the introduction.

In view of the uniformly insignificant correlations it may be assumed that phosphene threshold is not related to any of the visual functions included in the study. The lack of correlation with absolute threshold for light, on the basis of our previous argument, is interpreted to mean that the two thresholds reflect sensitivity of different structures, probably the photochemical receptors and the bipolar or ganglion cells of the retina. It seems that discrimination of phosphenes as light superimposed on the background illumination, involves different processes than the discrimination involved in equating one light intensity to another. The correlation with brightness discrimination is tenuous because of the low reliability of the measure. In the previous study with retardates, however, brightness discrimination was highly reliable but the correlation with phosphene threshold was of a similar order of magnitude.

To the extent that the phi-phenomenon is a measure of the movement-like component at threshold intensities of phosphenes, this component does not seem to be a major contributor to the threshold value.

The low correlation to CFF (also found in the previous retardate study) may be a result of the different nature of the two measures. CFF is primarily concerned with temporal factors by varying frequency, while the phosphene threshold is a measure of stimulus magnitude. Nevertheless, since the temporal resolving power of the visual system is an intimate function of the stimulus magnitude, one might expect a relationship between the two factors. The fact that no correlation is found may be interpreted as an indication that the two measures (phosphenes and CFF) differ in other ways. Previous studies have shown that fusion of electrical phosphenes is not found at frequencies well above those giving fusion of photic stimulation (2, 7). If CFF is a function of temporal resolution in the cortex (1), it is difficult to see why electrical phosphenes should not also fuse. Thus an alternative explanation for the lack of correlation would be that phosphenes result from stimulation of neural fibers in the retina, but the CFF is primarily a result of the inability of the photochemistry of the eye to respond to rapid flicker rate (8).

The lack of correlation with cutaneous thresholds indicates that the phosphene threshold does not reflect, in general, the electrical excitability of the organism. This lack of correlation, however, may not be conclusive. There are indications that the cutaneous thresholds are not very stable measures. Individual variability is higher for the cutaneous than for the phosphene thresholds and the reliability for cutaneous threshold at 20 cps is low even though significant at the five per cent level. This is even more apparent when each *S*'s threshold determinations obtained in a single session are scrutinized. While phosphene thresholds show minor variability, the cutaneous thresholds as a general rule vary considerably. Added to this is the report from several *S*s that the quality of the sensation at threshold intensity was not the same from determination to determination, or from day to day. It was also observed that there seemed to be a direct relation between amount of subcutaneous fat and cutaneous thresholds. It is thus indicated that there are many parameters of the skin thresholds that need to be investigated. In spite of the convenience of electrode attachment for cutaneous thresholds, this sensory modality may have been an unfortunate choice for investigating the problem of phosphene threshold as an indicator of general neural excitability.

In conclusion it may be stated that according to expectations, phosphene threshold is not correlated to visual functions involving adequate stimuli. These findings are interpreted to indicate that the phosphene threshold and the other visual functions reflect sensitivity of different structures. More specifically, the findings are consistent with the view that at threshold intensities the electrical stimulation of the eye affects neural tissue of the retina, probably the bipolar or ganglion cells. This does not exclude the possibility that at higher intensities several types of nerve fibers may be involved or that photochemical processes may be affected as a secondary manifestation. Our findings leave it an open question whether phosphene threshold is an indicator of neural sensitivity to electrical stimulation in general.

The significance for the organism of the sensitivity of the structure involved in the phosphene threshold remains obscure, even though some possibilities have been ruled out by this and previous studies. The most hopeful lead for understanding the functional characteristics reflected in the threshold seems to be the finding of elevated thresholds in retarded *S*s, and the relation to per cent alpha in retarded *S*s with clinical evidence of brain pathology. It is our intention to verify and extend this lead.

E. SUMMARY

In phosphene experiments an alternating or intermittent current is applied through surface electrodes across the eye. It is generally assumed that the electrical current is acting upon neural tissue, such as the bipolar or ganglion cells of the retina, and that the phosphene threshold reflects the sensitivity of these structures. On this basis it was predicted that phosphene threshold would not be highly correlated with various visual functions. In 14 normal adults, phosphene threshold at 20 and 50 cps was compared to absolute threshold for light, brightness discrimination, CFF and perception of apparent motion (phi-phenomenon). While test-retest reliability for most of these tests were high, the correlations with phosphene thresholds were not significant. This is regarded as support of the premises concerning phosphene threshold.

In addition, the question was asked if phosphene threshold could be re-

garded as a general indicator of neural excitability. For this purpose the cutaneous sensation threshold for electrical stimulation was determined for the forearm. Insignificant correlations were found, but since the cutaneous threshold appears to be a rather unstable measure, this finding may not be conclusive.

REFERENCES

1. BARTLEY, S. H. Central mechanisms of vision. *In* Field, J., Magoun, H. W., & Hall, V. E. (*Eds.*), *Handbook of Physiology*, Section I: *Neurophysiology*. Vol. I, Washington: *Amer. Physiol. Soc.*, 1959, Pp. 713-740.
2. BOUMAN, H. E. Rhythmical electrical stimulation of the retina and demonstration of an apparatus for rhythmical electrical stimulation. *Acta Brev. Néerl. Physiol.*, 1934, **4**, 114-115.
3. CLAUSEN, J. Visual Sensations (Phosphenes) Produced by AC Sine Wave Stimulation. Copenhagen: Ejnar Munksgaard, 1955.
4. CLAUSEN, J., & KARRER, R. Phosphene threshold as related to age and sex. *J. of Psychol.*, 1959, **47**, 189-198.
5. ———. Electrical sensitivity of the eye in the mentally retarded. *Train. Sch. Bull.*, 1961, **58**, 3-11.
6. ———. Aftereffect of electrical stimulation to the eye. *J. of Psychol.*, 1962, **53**, 105-118.
7. SCHWARZ, F. Quantitative Untersuchungen ueber die optische Wirkung sinusförmiger Wechselströme. *Z. Sinnesphysiol.*, 1940, **69**, 1-26.
8. SVAETICHIN, G. Receptor mechanisms for flicker and fusion. *Acta Physiol. Scand.*, 1956, **39**, Supplement 134, 49-54.

Department of Research
The Training School
Vineland, New Jersey

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 101-108.

PSYCHO-TECHNOLOGY IN HIGHER EDUCATION VERSUS PSYCHOLOGIZING*

College of Education, University of Arizona

SIDNEY L. PRESSEY

It has been estimated that within ten years college enrollments will double. And increased funds are being desperately sought to provide buildings and staff to handle this flood of students—in much the same old ways as formerly! All available funds will indeed be needed. But the prediction is ventured that sweeping changes in methods will be needed also, and that such changes as here are briefly suggested could convert a crisis into a break-through into far greater educational effectiveness than ever before, providing the nation with more and better-trained college graduates than would otherwise eventuate.

A. OVERLAP VERSUS OPPORTUNITY

A notable 1924 investigation of secondary school-college curricular overlap (6) found approximately a fifth of the first two years of college repeated work in secondary school; half the selections read in a college course in English literature might have been in a high school course in that subject; there might be a 70 per cent overlap of laboratory work in chemistry. Yet so little was the situation remedied that a 1952 inquiry (2) found similar duplication: a third of a large group of university students took American history although having just completed a course in that subject in secondary school. There was similar duplication in English literature. In science, almost half of the college students took the beginner's course they had already taken in high school and many so repeated two sciences, though grades earned indicated negligible gain from the repetition.

Largely from this last investigation did emerge the Advanced Placement Program of the College Entrance Examination Board. It presumes advanced secondary school courses, and so serves only the students from schools able to give such courses. Nevertheless in 1961 over 1000 schools prepared more than 13,000 candidates for advanced placement, in 600 colleges and universities. With what success? Harvard has led in support of the program (18). There, some 45 per cent of the '61 entrants were advanced placement candidates; of

---

* Received in the Editorial Office on August 26, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

these, 83 per cent were given advanced placement and a fourth of them in three or more subjects, thus making them eligible at entrance for sophomore standing. And records of those in previous years given advanced placement showed them doing well. Of 71 who had graduated, 55 had done so in three years and of these last, four-fifths with honors. It would seem well instanced that, in such an institution, the advanced placement program of the CEEB can cut down school-college overlap with desirable academic outcomes.

Without the presumption of advanced secondary school courses, which increase school-college overlap as a preparation to reduce it, is the long-established but too little-known procedure at the University of Buffalo (5). A high school student who believes (or whose teacher judges) that he might pass a course at that institution for credit by examination, may obtain an outline of the course with readings and sample examinations, and prepare by independent study with a modicum of help from his teacher. Repeated follow-ups have shown that students thus credited averaged better in later college courses than students of similar ability who took the prerequisite courses instead of eliminating them by examination, also saved an average of nearly half a year in college *and* more often continued into graduate or professional training.

Without any involvement of the high school is the provision in many institutions that a student may upon application, (*a*) take an anticipatory examination in any course which if passed gives credit, (*b*) take a placement examination most commonly in English composition at the beginning of the freshman year on which a very high score gives credit. By such provisions, about a fourth of the graduating students (a less proportion of all those entering) at Ohio State University have obtained some examination credit and a sixteenth the equivalent of a quarter's work or more (11). And again, follow-ups showed those with some such credit doing superior academic work both generally and in the subjects in which the credit had been obtained.

However, many students come from high schools not offering advanced work, do not know about applying for credit-yielding examinations, or find faculty members hostile to the idea. Further, very able youngsters may be ahead of what might be expected. So Learned and Wood (7) found almost 10 per cent of Pennsylvania high school seniors scoring above the median for college seniors, on their battery of tests supposedly covering the essentials of a college education. A group of gifted California high school seniors "surpassed the achievement of the average college senior on tests in the social sciences, humanities, and natural sciences" (10, p. 2). A procedure meeting these and other needs would seem to be the giving in freshman week (or at the beginning of each large freshman course) of placement tests to all students, with credit for scores above a determined point. Such an arrangement at Antioch College gave credit by examination to 64 per cent of the students entering in 1961; 23 per cent were given 15 hours or more. When such credit was obtained, students usually went at once into a "Level II Course," and there did on the average slightly better than those who took the Level I course at Antioch (3).

Harvard, Buffalo, Ohio State and Antioch are institutions each differing from the others in many respects, among them their procedures in reducing school-college overlap. But all have found credit by examination substantially to reduce it, with ensuing good academic work and increase in numbers continuing into advanced training. Surely such advantages should not be limited to students from high schools offering special relevant advanced courses, nor require special permission to attempt, nor entail a special burden on faculty members or departments willing to give such examinations. Rather (somewhat as at Antioch) all entering students should be thus appraised, with credit given above a carefully determined cut-off point and with probable sectioning according to score below that level. The possibility of such crediting should stimulate both high school teachers and students. It seems a reasonable inference that a third of all freshmen might obtain some credit and a tenth so dispose of at least half their first year requirements. Such saving in number of freshman sections, and in student time, seem not negligible.

B. ALL-YEAR SCHOOL?

Full utilization of an educational plant obviously requires operation throughout the year. But what may be the effects on students of all-year schooling? The many so attending during and immediately after the second world war gave the writer exceptional opportunities to investigate the matter (11, pp. 114-120). Careful comparisons of those students with others not in school summers yielded no consistent indications that four-quarter attendance led to poorer academic work. Nor did inquiry forms and individual interviews indicate undue fatigue or ill health, independent of war-time stresses, or concomitant part-time employment—after all, a four-quarter year still leaves a total of about six weeks vacation in September and between other quarters. There was somewhat less participation in undergraduate social life. But this is traditionally organized on a September-June school year (and four of them) to accumulate "seniority" for status and honors. It should be possible to organize social life on a quarter or trimester basis so as to avoid such difficulties.

Desirable work or travel experience might be missed with a four-quarter year, also escape from the too-close family ties of some commuting students. But for some it was found possible to arrange a quarter or two of relevant employment (given some university credit) which did meet these last mentioned needs, and could be of great educative value as seems now to be increasingly recognized (19). It was concluded that comprehensive all-year plans, which included one or more quarters or trimesters of work or travel or off-campus projects as in New York or Washington, all tied in with on-campus studies, might usually be better than the usual school year with summers left to the student's own resources. And the student might finish an undergraduate program in three calendar years instead of four, thus being able to start advanced training or career sooner.

C. Independent Study and Teaching Machines

Over twenty-five years ago, an independent study plan for superior students was reported (17) involving no class meetings for as long as two weeks, with records in subject-matter equal to paired cases in regular classes, and gains in capacity for self-directed work. Recently with the support of the Ford Foundation (1) various types of independent study have been tried in a variety of institutions, with similar favorable results. In high schools it is being urged that independent study involve some 40 per cent of the student's time (16). It seems well evident that, with appropriate planning, "students are able to learn as well with much less class time than has customarily been required of them" (1).

Such savings of class time are possible with conventional instructional materials. Use of teaching machines permits yet more saving. Not only may relatively routine teaching tasks be thus managed, but special "programmings" may take the place of textbook or other conventional matter. As the writer has recently stressed (12), textbook and other matter may be kept, but simple pages of supplemental material (perhaps prepared by the instructor) may be used with them and made "auto-instructional" by such a simple device as a "chemo-card" on which, in the appropriate space, the student checks what he thinks is the correct answer to each objective question on the supplemental page; if he *is* right, his check-mark instantly changes color; if it does not change he knows he is wrong, and continues until he finds the right answer. With superior students, an auto-instructional laboratory using such materials, where each works as he wishes and consults with the assistant in charge when desired, may replace formal instruction and permit finishing a course in the first six weeks of a quarter. Several investigators have reported success with such procedures (9, pp. 82-8, 173-85). Able students thus not only save time, but can go through more courses in a given time; the writer's investigations have included students who completed, with honors, a four-year undergraduate program in two calendar years.

D. NET BENEFITS OR RISKS?

With such rapid progress, are subtle educational values lost, and is the graduate often immature? A follow-up ten years after graduation (at a median age of 20.9) of 104 students who had completed an undergraduate four-year program in three years or less and some in two, is here relevant (4). Students paired with them as to age and ability at entrance and type of curriculum, but who took the conventional four years, had a slightly *lower* undergraduate academic record, and neither earned money (summer or part-time while in school) nor participated in student activities very much more. In the ten years following graduation, half as many of the paired students had earned further degrees, while approximately the same number had married and had participated in community activities. Terman (15) similarly found his most accelerated early-graduating gifted cases doing best both in school and after. Women graduating from a teachers college at the age of 20 or younger have been found more successful in career than cases similarly paired with them but graduated at 22; and men who were graduated young from an Eastern arts college were more successful in careers than those graduating older (11, pp. 67-73). In short, many if not most college undergraduates are now capable of more and more intensive work than most curricula ask of them, are nowadays fairly well "socialized" in high school, can do with a little less "college life," and are often desirous to be on into advanced training or career.

Rather than graduating too young, the danger for many may well be that they are kept in school too long, since some 70 per cent of the graduates of many colleges go on into some form of advanced training. While the median age of receiving the doctorate in most fields is now into the thirties, the presidents of national professional organizations in those fields received the doctorate about five years sooner (13). A wide range of evidence supports the implication (massively documented by Lehman) that greatest creativity comes in the younger adult years, into which mere "getting ready" should not unduly extend. That students are becoming increasingly dissatisfied with the postponement of realization of adult status is presumably indicated by this further finding: a sample of senior men in the 1961 autumn quarter of the Ohio State University (which records marital status) showed

32 per cent of them already married. Earlier completion of undergraduate education, by such means as proposed in this paper, might well be desirable sociologically and economically as well as in terms of earlier and more productive career.

E. ACTION RESEARCH VERSUS PSYCHOLOGIZING

But if there are to be such radical innovations as general crediting by examination at entrance, all-year school, and wide-spread use of independent study often with auto-instruction, how may such innovations be appraised and guided so that standards are maintained? With only two or three big departments (and the administration) willing to experiment, much can be done. So a pretest can be given at the beginning of a freshman sequence, those scoring in the top 15 per cent can be given credit for the first course and started in the second, and their record compared with the remainder required to take the first course, also the top 15 per cent the previous year. In the first six weeks of a five-hour course, four sections may meet only two hours, the remaining work being done in independent study, and the practice continued, modified, or dropped according to results on a mid-term. Various forms of auto-instruction may be tried the same way. With the cooperation of the student personnel office, all year school and field work may be tried out. It can be done; a large part of the writer's research for twenty-five years was of this sort.

A very recent thousand-page volume full of valuable information derived largely from subsidized research (14) is here in contrast. Three chapters totalling 89 pages deal with entering students with much emphasis on their diversity but *not* on ways of dealing with that diversity, as for instance the experiments of the Fund for the Advancement of Education on early entrance and advanced credit, and other work of this general type mentioned earlier in this paper. The index includes trivium and quadrivium but not trimester or quarter plans or all-year school. A graph shows the supposed development of the Ego and Impulse but none exhibit any of Lehman's monumental work on age and achievement, and his name is not in the author index. A chapter on dropouts does not mention the substantial evidence that acceleration is associated with their reduction. Interesting chapters describe "Student Peer-group Influence" and "Student Culture at Vassar" but with no reports of experiments attempting constructively to influence student groups. Yet that can be done, as instanced by several investigations under the writer's direction (see for instance Pressey, S. L., and Hanna, D. C. The class as a psycho-sociological unit, *J. of Psychol.*, 1943, **16**, 11-19). Faced with the rising tidal wave of students, and the nation's critical need for trained manpower, there is surely need not only to gather data about problems of higher education, but also to experiment vigorously with ways of dealing with them.

F. SUMMARY

This paper argues that the coming flood of students and present need for trained manpower should not be handled simply by spending more money to build more buildings and hire more staff to do the same old things in the same old way. The prediction is ventured that educational efficiency could be doubled by research-validated use of such means as: (*a*) credit by examination especially at entrance; (*b*) all-year school; (*c*) credited field work; and (*d*) independent study often aided by use of auto-instruction. Further, students could thus save a year, and, instead of graduating when too "immature," would give more opportunity to obtain advanced training, thus entering their careers earlier in their years of greatest potential accomplishment and greatest usefulness to the country.

REFERENCES

1. BASKIN, S. Quest for Quality: Some Models and Means. New Dimensions in in Higher Education No. 7. Washington: U. S. Office of Education, 1960. Pp. 18.
2. BLACKMER, A. R. (chairman) General Education in School and College. Cambridge, Mass.: Harvard Univ. Press, 1952. Pp. 142.
3. CHURCHILL, R., & JOHN, W. Some evidence of the effectiveness of the achievement examinations in the general education program. Mimeographed report, Antioch College, September, 1960, also personal letter from Dr. Clarence Leuba of the Antioch faculty.
4. FLESHER, M., & PRESSEY, S. L. War-time accelerates ten years after. *J. Educ. Psychol.,* 1955, **46**, 228-238.
5. JONES, E. S., & ORTNER, G. College credit by examination. Univ. Buffalo Studies **21**, No. 3, January 1954. Pp. 123-201.
6. KOOS, L. V. The Junior College. Minneapolis: Res. Publications Univ. Minnesota, 1924.
7. LEARNED, W. S., & WOOD, B. D. The Student and His Knowledge. New York: Carnegie Foundation for the Advancement of Teaching, 1938. Pp. 405.
8. LEHMAN, H. C. Age and Achievement. Princeton, N.J.: Princeton Univ. Press, 1953. Pp. 359.
9. LUMSDAINE, A. A., & GLASER, R. Teaching Machines and Programmed Learning. Washington: National Education Association, 1960. Pp. 724.
10. MARTINSON, R. Educational Programs for Gifted Pupils. Sacramento: State Dept. Educ., 1961. Pp. 274.
11. PRESSEY, S. L. Educational Acceleration: Appraisals and Basic Problems. Columbus, Ohio: Ohio State Univ. Press, 1949. Pp. 153.
12. ———. Basic unresolved teaching machine problems. *Theory into Practice,* 1962, **1**, 30-37.

13. ———. Age and the doctorate—then and now. *J. Higher Educ.*, 1962, **33**, 153-160.
14. SANFORD, N., & OTHERS. The American College. New York: Wiley, 1962. Pp. 1084.
15. TERMAN, L. M., & ODEN, M. The Gifted Child Grows Up. Stanford, Calif.: Stanford Univ. Press, 1947.
16. TRUMP, J. L. Images of the Future: A New Approach to the Secondary School. Washington: National Educ. Assoc., 1959.
17. UMSTATTD, J. G. An independent study plan. *J. Higher Educ.*, 1935, **6**, 143-8, 240-43, 364-7.
18. WILCOX, E. T. A Report to the Faculty of Arts and Sciences on the Program of Advanced Standing, Cambridge: Harvard Univ., 1962. Pp. 16.
19. WILSON, J. W., & LYONS, E. H. Work-Study Programs. New York: Harpers, 1961.

College of Education
University of Arizona
Tucson, Arizona

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 109-113.

ON THE ELIMINATION OF ARTIFACTS IN GSR CONDITIONING*

Payne Whitney Psychiatric Clinic, New York Hospital, and Department of Psychology, Hunter College

SAM J. KORN AND LIVINGSTON WELCH

A. INTRODUCTION

The central issue of this study is the examination of GSR adaptation in two specific programs of stimulus presentation. In the pursuit of more adequate conditioned response methodology, a technique was recently introduced (1) that preserves the advantages of GSR and its sensitivity, and at the same time strikingly reduces the artifacts of pseudo-conditioning that are ordinarily compounded with conditioned GSR data. In this earlier paper, it was demonstrated that adaptation achieved through conventional techniques, such as the repeated unreinforced presentation of the CS, is incomplete. "Adaptation" thus achieved is only transitory, and the post-conditioning responses, as has often been argued, may very well represent the reoccurrence of preconditioned, or orienting, responses to the CS.

In the previous paper, a technique for adapting GSR was described that involves the continuous presentation of one or another of four colored signal lights. The order of presentation of the four colors was an unsystematic one, and as one light went off, the next light immediately went on. Thus, there was no inter-stimulus interval; one of the four lights was always being projected. This technique may be referred to as the Continuous Series Method in contrast with the more usual techniques that may be characterized as the Intermittent Stimulus Method. It was clearly demonstrated that adaptation of GSR obtained through the Continuous Series Method was quickly attained and reliably maintained for at least 45 presentations of the arbitrarily preselected "critical stimulus"—an amber signal light. This was *not* the case for adaptation achieved through the Intermittent Stimulus Method; GSR to the same amber light, aperodically presented and with no imposed reinforcement, persisted or recurred in an irregular manner for 25 presentations. In both methods, the duration of the adaptation trials was approximately 25 minutes.

---

* Received in the Editorial Office on August 26, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

A further question that arises, however, concerns the possibility that in the latter technique, the conventional method of achieving adaptation, it is the aperiodic nature of stimulation that prevents adaptation of GSR. That is, in the Intermittent Stimulus Method, it is the change from the *absence* of a focal stimulus to the *imposition* of the signal stimulus, the CS, that serves as the continuing source of excitation and maintains the orienting response and its various components. This experiment is designed to examine the question by comparing the results of GSR adaptation obtained previously by the Continuous Series Method with adaptation achieved by a *Continuous* rather than Intermittent Stimulus Method. The stimulus in this situation will be the same amber light, but its presentation will be uninterrupted and unreinforced.

B. PROCEDURE

*S*s were 14 female college students who were enrolled in psychology courses at Hunter College. The details of apparatus and equipment were presented previously (1).

Individual *S*s were seated in a small, dimly illuminated room. Palmar electrodes were affixed and the *S*s were instructed to pay attention to the stimuli that were going to be presented. After a few minutes, a small aperture (1.6 inches, covered with ground glass) on the wall opposite the *S* was illuminated by an amber light. This signal remained unchanged for a period of at least 25 minutes. During this interval, GSR was continuously recorded.

Because, in this situation, stimulation was continuous rather than episodic, the conventional GSR measurement technique required some modification. Resistance changes to each presentation of the CS, obviously, could not be employed. In the previous experiment (1) the duration of each of the four stimuli alternately and continuously presented was 7.5 seconds. In order to facilitate comparisons, each *S*'s GSR record in the present study was subdivided into a continuous series of 7.5 second intervals. Starting with the third such interval after the amber light was first presented, every fourth 7.5 second interval was designated as a "response interval" for the measurement of GSR changes. This was done to match the rate of appearance of the "critical" stimulus in the previous experiment. Thus, in both of these studies, 45 recorded responses to the *same amber light signal* were recorded for each *S*. Response magnitude was computed, in percentages, by comparing *S*'s resistance level immediately prior to the "response interval" and the peak resistance decrease within the 7.5 second duration of this interval.

C. RESULTS AND DISCUSSION

The results indicate that adaptation of GSR to the same stimulus, although continuously presented and unreinforced, was not completely achieved. For example, the median number of responses in the series of 45 "response intervals" was 14.5. In fact, a total of 10 or more responses occurred in 10 of the fourteen *S*s. Although these numbers appear to be small, in the Continuous Series Method, by comparison, only four of the fourteen *S*s yielded the same number of responses. The mean number of galvanic skin responses in the Continuous Stimulus Method was 13.2, with an *SD* of 3.8. This was significantly greater than the mean of 8.5 responses and an *SD* of 2.6 obtained in the earlier experiment. The difference between the mean number of responses in the two methods was statistically significant beyond the .01 level ($t = 3.67$).

Not only was the number of responses greater, but the *magnitude of GSR* was also greater in the stimulus situation that involved the continuous single stimulus. This comparison is summarized in Table 1. For example, approximately 23 per cent of the responses in the Continuous Stimulus Method were over 6 per cent, while in the Continuous Series Method, only about 7 per cent of the responses were over this response level. The chi square of the presented contingency table is significant beyond the .01 level of confidence.

TABLE 1
FREQUENCY OF RESPONSES (GSR) OF DIFFERENT MAGNITUDES IN TWO TYPES OF ADAPTATION PROCEDURES

	Magnitude of response in per cent					
	0–2	2.1–4	4.1–6	6.1–8	above 8	Total
Continuous stimulus	71	44	28	18	24	185
Continuous series	90	17	4	6	2	119

The difference between these two types of adaptation techniques is not only apparent in the over-all trends of the responses. Even when only the final 10 "response intervals" or stimulus presentations are considered, the advantages of the Continuous Series Method are clear. The number and magnitude of responses to the last ten trials, in both methods, are presented in Table 2. Both the number and magnitude of responses are greater in the Continuous Stimulus Method—even after 35 "response intervals" and 18-20 minutes of stimulation without reinforcement.

Thus, GSR adaptation, the absence of responses to an unreinforced visual stimulus, as was previously demonstrated, is most completely achieved when

TABLE 2
DISTRIBUTION OF MAGNITUDE OF RESPONSES (GSR)
IN LAST 10 OF 45 "ADAPTATION" TRIALS

Response magnitude	Continuous stimulus	Continuous series
0.1–0.5 per cent	0	0
0.6–1.0	0	8
1.1–2.0	7	3
2.1–3.0	4	3
3.1–4.0	2	1
4.1–6.0	4	2
6.1–8.0	4	2
above 8.0	7	0
	25	18

the adaptation procedure includes the continuous presentation of one or another of a series of similar stimuli. A single stimulus, on the other hand, whether presented intermittently or without interruption, does not induce the same level of adaptation of the GSR. When only a single stimulus is presented, adaptation of GSR is not reliably maintained. In the conventional conditioned response procedures, this reappearance of the orienting responses would be hopelessly confused with conditioned responses per se.

It is apparent from these results that GSR adaptation achieved through the Continuous Series Method is unique and strikingly superior. A complete accounting of the precise conditions that underlie this superiority has not, as yet, been made. The effects of variation of stimulus intensity, rate of alternation, etc. within different sensory modalities must be determined. However, it is quite clear that the troublesome persistence of GSR variation in so-called "neutral" stimulus situations is virtually absent in this situation. The flow of associations, with the concomitants of emotional arousal and GSR reaction, so often found in conventional adaptation techniques, does not occur in the Continuous Series Method. This method appears to provide a *continuous and stable* state of orientation that does not easily yield to intrusion and GSR fluctuation.

At this point, the Continuous Series Method represents, essentially, a practical procedural control. It will, however, be of no avail if GSR apparatus and recording equipment are themselves erratic and subject to error. This cannot be overemphasized. The methodological control of GSR, with its sensitivity and apparent irregularity, can only follow after apparatus features are absolutely accurate and reliable. Within this framework, the data reported strengthen the position taken in our previous paper regarding the problem of adaptation of GSR; the vagaries of pseudoconditioning *can* be minimized and GSR conditioning may be more reliably pursued.

D. SUMMARY

Adaptation of GSR to a continuously presented visual stimulus (Continuous Stimulus Method) was compared with results obtained in an earlier study in which *S*s were continuously presented with one or another of four signal lights (Continuous Series Method). In each of these studies, 14 female *S*s were employed and the duration of stimulation was approximately 25 minutes. Although no reinforcement was imposed, the number of responses and magnitude of GSR were greater in the Continuous Stimulus Method than in the Continuous Series Method. In the latter, adaptation was more quickly accomplished and persisted more reliably. Thus the advantages of the adaptation procedure involving the continuous presentation of a succession of similar stimuli are unique and the implications for CR procedure are reaffirmed.

REFERENCE

1. KORN, S. J., & WELCH, L. Conditioned responses: fact and artifact. *J. of Psychol.*, 1962, **53**, 301-309.

Payne Whitney Psychiatric Clinic
525 East 68th Street
New York 21, New York

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 115-119.

ACADEMIC ACHIEVEMENT CORRELATES OF ACHIEVEMENT AND AFFILIATION MOTIVATIONS*[1]

University of Buffalo

LAWRENCE W. LITTIG AND CONSTANTINE A. YERACARIS

A. PURPOSE

In a recent analysis of high school achievement, Rosen (4) has reported that achievement motivation (*n* Achievement) is related to course grades. High *n* Achievement male high school subjects (*S*s) obtained significantly higher course grades than their low *n* Achievement counterparts.

Veroff, Atkinson, Feld, and Gurin (6) have also reported finding relationships between *n* Achievement and educational level of achievement. On the basis of an analysis of materials collected from a national sample they report that the percentage of *S*s with college education, regardless of sex, who have high *n* Achievement scores is greater than the percentage of *S*s who have only a grade school education. In addition, they report that affiliation motivation (*n* Affiliation) and power motivation (*n* Power) show a similar relationship with education among women but not among men.

The present paper reports further relationships between education and motivation and speculates about the possible meaning of these relationships taken together with those cited above and with certain findings about motivation and early childhood experiences (1, 7).

B. METHOD

1. *Subjects*

The data resulted from a secondary analysis of an investigation of the effects of educational television on the attitudes of a community's members toward education. A probability cluster multi-stage sample of 400 heads-of-households (men or unmarried women) or the wives of heads-of-households was obtained. One hundred and ninety men and 210 women were interviewed. Data relevant to this report were not obtained from four women reducing the female *n* to 206.

---

* Received in the Editorial Office on August 27, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Financial support for this study was received from the Ford Foundation and the University of Buffalo Graduate School.

All *S*s were interviewed in their homes by trained interviewers who administered a 90-item pretested schedule to assess attitudes toward education. At the end of the interview *S* was asked his highest level of academic achievement. Prior to responding to the questionnaire, however, *S* was requested to tell stories to a six picture thematic apperception series. The six pictures had been used previously in the Veroff, *et al.* national survey (6). The stories obtained from *S*s were coded for *n* Achievement by means of the content analysis technique described by McClelland, Atkinson, Clark, and Lowell (3) and for *n* Affiliation according to the method described by Heyns, Veroff, and Atkinson (2). The coding was done by four expert coders who were trained using the manuals for *n* Achievement and *n* Affiliation prepared by Smith and Feld (5). Distributions of *n* Achievement and *n* Affiliation scores were prepared within interviewer-coder pairs to reduce any systematic bias and *S*s were classified as being high on these motivation indices if their scores were above their respective groups' medians and low if their scores were below their respective groups' medians. *S*s were separately grouped for *n* Achievement and for *n* Affiliation for purposes of analyzing their relationships to academic achievement. Academic achievement was separated into four levels which appeared, *a priori,* to be most significant. These were: less than high school graduation; high school graduation; some college; and college graduation and beyond.

C. RESULTS

Data from men and women are treated separately.

1. *Achievement Motivation and Academic Achievement*

Achievement motivation is related to academic achievement for men ($X^2 = 11.20$, $df = 3$, $P < .02$) but not for women ($X^2 = 4.92$, $df = 3$, $P > .10$) though the trend for women is similar to the pattern obtained for men (Table 1). It is apparent in Table 1 that the relationship between *n* Achievement and academic achievement is essentially due to the distributions of *S*s without a high school diploma and *S*s with a college degree or more. In the former group low *n* Achievement *S*s are in the majority (36 per cent high *n* Achievement vs. 52 per cent low *n* Achievement) whereas in the latter group high *n* Achievement *S*s appear with much higher frequency (33 per cent high *n* Achievement vs. 14 per cent low *n* Achievement).

2. *Affiliation Motivation and Academic Achievement*

No statistically significant relations between *n* Affiliation and academic achievement are present for either men or women (Table 2).

TABLE 1
ACADEMIC ACHIEVEMENT CORRELATES OF *n* ACHIEVEMENT

	Men				Women			
	High *n* Ach		Low *n* Ach		High *n* Ach		Low *n* Ach	
Academic achievement	N	%	N	%	N	%	N	%
< High school graduation	34	36	50	52	31	31	47	45
High school graduation	18	19	23	24	27	27	25	24
Some college	11	12	10	10	27	27	23	22
College graduation	31	33	13	14	16	15	10	9
Total	94	100	96	100	101	100	105	100

X^2 (men) $= 11.20$, $df = 3$, $P < .02$.
X^2 (women) $= 4.92$, $df = 3$, $P > .10$.

D. DISCUSSION

The results lend themselves to several plausible inferences. First, a predictive validation of thematic apperception as a method of assessing motivation is provided by the positive relationship between *n* Achievement and academic attainment and the absence of a relationship between *n* Affiliation and academic achievement, at least for men. This is as it should be if the *n* Achievement coding system is tapping a disposition to strive for success in competitive achievement situations (the achievement motive) and the *n* Affiliation coding system is tapping a disposition to gain satisfaction from positive social responses from others (the affiliation motive). Further, this positive relationship between *n* Achievement and education reported here is

TABLE 2
ACADEMIC ACHIEVEMENT CORRELATES OF *n* AFFILIATION

	Men				Women			
	High *n* Aff		Low *n* Aff		High *n* Aff		Low *n* Aff	
Academic achievement	N	%	N	%	N	%	N	%
< High school graduation	41	44	43	45	37	35	41	41
High school graduation	25	26	16	17	25	24	27	27
Some college	10	11	11	11	31	30	19	19
College graduation	18	19	26	27	12	11	14	13
Total	94	100	96	100	105	100	101	100

X^2 (men) $= 3.61$, $df = 3$, $P > .20$.
X^2 (women) $= 3.11$, $df = 3$, $P > .20$.

consistent with the findings of Rosen (4) and Veroff, *et al.* (6) regarding *n* Achievement and education, and the absence of a relationship between *n* Affiliation and education for men is in line with the Veroff, *et al.* result.

However, it may be asked if the *n* Achievement findings are not simply the result of greater verbal ability of educated *S*s. If this be the appropriate explanation of the findings, there is every reason to believe that *n* Affiliation scores would also be raised by increased verbal fluency among educated *S*s. Veroff, *et al.* (6) have reported that female *S*s show a tendency to be higher on *n* Affiliation with increasing level of education but this is the only reported case of such a finding. At least for male *S*s the evidence is consistent. If presently available evidence reduces the probability of verbal ability solely accounting for the relationship between *n* Achievement and academic achievement the problem of causality still remains unsolved. Is *n* Achievement one of the factors which influence success in academic affairs or do educational experiences tend to increase the probability that *S* will verbalize achievement themes? The presently available data do little more than emphasize our lack of knowledge about this problem but, taken with recent evidence presented by Winterbottom (7) and Feld (1) that indicates that the achievement motive is acquired early in life as a result of parental attitudes about achievement and independence training, it seems plausible at this time to argue that *n* Achievement may be a prior condition to academic achievement for men, at least, rather than a result of later learning in school.

E. SUMMARY

Academic achievement level, *n* Achievement scores, and *n* Affiliation scores were obtained from a probability sample of 190 men and 206 women residing in a small community selected for a study of educational television. It was found that *n* Achievement was positively related to academic achievement among men but not among women. Need Affiliation was not related to academic achievement for either sex. Implications were drawn from these findings regarding: (*a*) the validity of the thematic apperceptive measure of *n* Achievement; and (*b*) the relevance of *n* Achievement to academic success.

REFERENCES

1. FELD, S. C. Studies in the origins of achievement strivings. Unpublished doctoral dissertation. Univ. Michigan, 1959.
2. HEYNS, R. W., VEROFF, J., & ATKINSON, J. W. A scoring manual for the affiliation motive. *In* Atkinson, J. W. (*Ed.*), *Motives in Fantasy, Action, and Society*. Princeton, N. J.: D. Van Nostrand, 1958.
3. MCCLELLAND, D. C., ATKINSON, J. W., CLARK, R. A., & LOWELL, E. L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

4. ROSEN, B. C. The achievement syndrome: A psychocultural dimension of social stratification. *Amer. Sociol. Rev.*, 1956, **21**, 203-211.
5. SMITH, C. P., & FELD, S. C. How to learn the method of content analysis for *n* Achievement, *n* Affiliation, and *n* Power. *In* Atkinson, J. W. (*Ed.*), *Motives in Fantasy, Action, and Society*. Princeton, N. J.: D. Van Nostrand, 1958.
6. VEROFF, J., ATKINSON, J. W., FELD, S. C., & GURIN, G. The use of thematic apperception to assess motivation in a nationwide survey. *Psychol. Monogr.*, 1960, **74**, No. 12 (Whole No. 499).
7. WINTERBOTTOM, M. R. The relation of need for achievement to learning experiences in independence and mastery. *In* Atkinson, J. W. (*Ed.*) *Motives in Fantasy, Action, and Society*. Princeton, N. J.: D. Van Nostrand, 1958.

National Science Foundation
Washington 25, D. C.

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 121-143.

SOME RELATIONS BETWEEN SENSORY END RESULTS AND NEURAL ACTIVITY IN THE OPTIC PATHWAY*[1]

Department of Psychology, Michigan State University

S. HOWARD BARTLEY AND THOMAS M. NELSON

A. ORIENTATION

The purpose of this paper is to account for certain major aspects of vision such as brightness, by presenting material showing how the optic pathway functions. This involves a description of the relation of various neural response processes evoked by patterns of photic impingement. The intent is to present a connected and coherent picture of the role of temporal factors in producing sensory outcome.

Much individual experimentation in neurophysiology, as in other disciplines, is not designed to answer general questions dictated by systematic considerations. Presentday information leaves certain types of gaps in a connected understanding of just what takes place in all parts of the pathway to selected sorts of input to the eye. Some information on hand is much more effective than others in making a start in developing this understanding. It is on the basis of this relative effectiveness that we shall select from the available facts.

1. *Biological Mechanism and Sensory End Result*

In visual science one is interested in understanding how the relation of external event to sensory end result is possible. The whole activity from periphery to center is a succession of inputs and consequent outputs from several waystations, each producing modifications until the sensory end result eventuates. Explanation therefore demands an account of the full sequence of relevant biological events.

It has long been a rule, as well as a convenience, to begin at the sensory periphery rather than at the center when analyzing features influencing the overall behavior of an organism. One begins by studying the sense organ, the first mechanism for discriminating external events, rather than beginning with the central nervous system. However, what is discovered and inferred

---

* Received in the Editorial Office on August 27, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

about the way the periphery acts is not a complete explanation of the mechanisms underlying sensory end results. Knowledge of the periphery cannot take the place of directly observed central nervous activity in interpreting how the central nervous system acts (12).

This obvious fact has sometimes been forgotten. There was, for example, a rather prominent era in which the quantitative features of sensory end results were attributed solely to photochemical events in visual receptors. The leaders in this interpretation argued that the essential nature of the sensory end result was fundamentally determined by the photochemistry of the rods and cones (41, 42). During the heyday of such interpretation (the 20's, 30's and even the early 40's) it was thought improper to assign any role to the neuroretina in modifying the quantitative output of receptor photochemistry and therefore participating in determining sensory end results. The neuroretina was customarily treated as simply part of the neural conduction system.

Since that time a great many facts of retinal neurophysiology have accumulated, and the behavior of receptors as neural units has become better known. But today the same tendency that we have just referred to is being duplicated. Now there are those who tend to regard the neuroretina as the prime determiner of the sensory end result, leaving the central end of the neural pathway only a secondary role (26).

2. *Clear Terminology Needed*

Before the facts are stated, assimilated, and interpreted, the character of descriptive and theoretical language used by those studying vision must be settled upon. Confusion is unavoidable when language is carelessly used. Let us begin by disentangling two closely related orders of activity, the energistic and neurophysiological.

"Stimulus" is a loaded term. It implies activation. Nevertheless it is frequently applied to situations in which the energy applied is ineffective. This produces a paradox since the "stimuli" do not stimulate. To obviate this misleading and inappropriate indication, we will apply the term *stimulus only* to circumstances where external energy applied to the receptors or to the optic nerve produces a measurable effect in the kind of processes dealt with in the discipline involved. In physiology, the effect is in terms of some body process; in psychology it is in personalistic terms, such as sensory end result. It is only from consideration of biological results and not by the measurement of the impinging energies that stimulation can be said to occur or that a "stimulus" has been applied. We shall, therefore, use the word *impingement* when referring to energies per se. An impingement is simply applied energy. It may or may not produce phenomena in the domain being dealt with (10).

The necessity of clearly separating terms referring to (*a*) impingement and (*b*) peripheral or optic pathway phenomena from (*c*) sensory end results is seldom more desirable or difficult than in dealing with intermittent stimulation. Since the situations we shall deal with systematically involve some combination of these categories, our language must give clear evidence of which categories are involved.

One group of studies is customarily called "flicker experiments" because they involve use of a range of intermittent photic impingements wherein flicker is sometimes the sensory end result. These studies are poorly named. With frequencies of a high enough order, uniform continuous illumination rather than flicker is seen. Transferring the term "flicker," which is clearly a response term, to the impingement and calling impingement rate "flicker rate" is needlessly confusing. Bartley (4) and Segal (52) have independently shown that the relation between photic pulse rate and flicker rate is not a one-to-one affair. At low intensities, for example, an impingement intermittency of 10 cps will yield flicker at about twice that rate, while an intermittency of 50 cps at an intensity 1000 times greater will be seen as flicker somewhere in the neighborhood of half the impingement rate. In the present paper we shall use *flicker* to refer only to the brightness undulations actually seen. And we shall call each single brief impingement a *photic pulse* or just simply a *pulse* and retain the term *flash* as a label for what may be seen in such cases.

Another group of "flicker" experiments is so-called because the rate of photic intermittency produces neurophysiological records showing a *neural response rate* (NRR) in phase with the photic pulses (29, 30). A further complication is the fact that, under many conditions, each *neural response* (NR) is made up of a group of impulses ("spikes"). A distinction seems desirable between neural response (NR) as a total reaction to a single impingement and *discharge rate* (DR), the frequency of impulses within such a response. (See Figure 1.)

Neural activity is not only erroneously identified as flicker but also as *critical flicker frequency* or CFF, the pulse rate just eliminating flicker (31). The records show that neural responses undergo changes from: rapid volleys of spikes, each response separated by an interval of slow discharge; to volleys without intervals of slow discharge; to shortened volleys; to responses consisting of single spikes; and to sporadic responses not elicited by successive photic pulses. This allows for certain correspondences between neural dis-

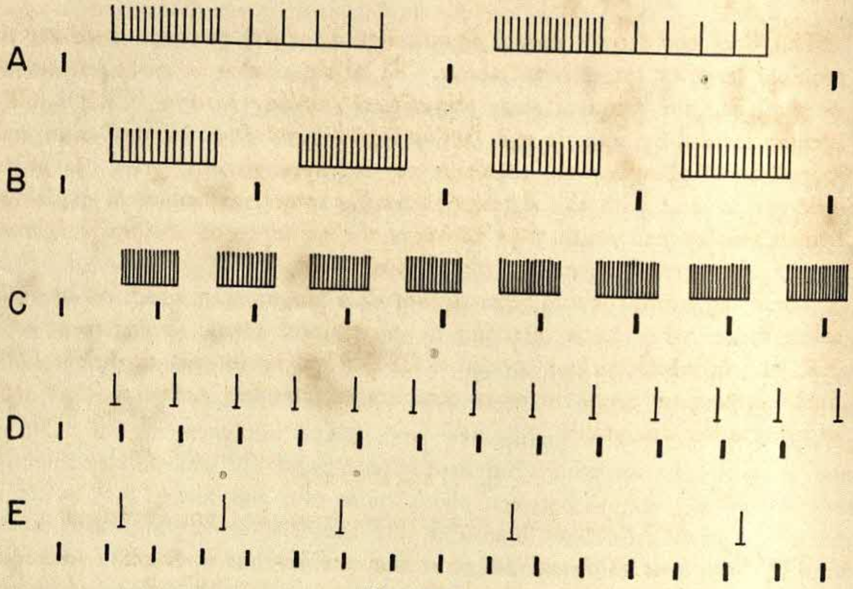


FIGURE 1

Schema showing relation between photic pulse rate, neural response, and the discharge rate (DR) within response. Upper notation in each line, responses; lower, photic pulse rates. A. Slow rate, each response is a burst of spikes followed by a slower discharge; B. Faster rate, response only a burst; C. Still faster rate, shorter bursts but DR higher; D. Still faster, each response is a single spike. It is supposed that this represents the neural discharge underlying CFF in the human, and is called the critical response pattern (CRP). E. Some elements discharge irregularly at this rate or they may fail to discharge at all.

charge and sensory end result to be inferred since there is also a series of steps or changes in the sensory end results. (See Figure 1.) The neural response rate at which uniform brightness is just produced in humans can be determined by noting the rate at which *S*s report uniform brightness. Brightness, it must be remembered, is a sensory phenomenon, not neural. A great amount of information in accounting for vision is obtained from experiments on subhuman species; and the neural records from both species obtained under comparable impingement conditions are quite similar. Imputing sensory experience to lower species in order to understand necessary interrelations is entirely gratuitous. It is well enough to recognize that as rate of impingement intermittency is increased, certain critical points in neural response are reached both in human and subhuman species. But, since we are using the subhuman species only to gain information about body mechanisms so as to account for sensory end results, the relating of body process to sensory end result needs to enter only in dealing with humans.

This does not deny a special significance to neural responses reaching the terminal stage or stages listed above. To be clear what is being referred to, we shall call the terminal stage the *critical response pattern* (CRP). CRP is characterized by a single spike following each pulse; or in some cases, only sporadically following the members of the impingement series. It is far different to deal with this pattern as having some significance in explaining human sensory end result than to imply the existence of uniform brightness or any other sensory experience in the species studied.

These suggestions would seem to provide a language in which no one talks about flicker when he is referring to intermittent photic impingement or to NR, and in which no one speaks of CFF when he intends to denote CRP. Such conventions seem no more complicated, nor less necessary than those of many other disciplines.

B. PROCEDURES

1. *Two Main Kinds of Recording from the Optic Pathway*

The two general forms of recording are the *microelectrode* recording technique which is used to determine the nature of the activity of specific elements, and *mass* recording to picture the overall activity of some segment of the pathway. Microelectrode recording is highly analytical, but suffers at times from not giving an integrated view of the functioning of a given segment. It is a sampling technique and conclusions regarding what the retina-as-a-whole or some other tissue mass does are arrived at only by statistical inference. It should be mentioned here that the kind of connected picture we are attempting to achieve has had to rely heavily for its connectedness upon information gained by mass recording.

Records from the eye may be obtained by either microelectrodes or mass-recording electrodes. The former are put into the retina itself and may pick up either receptor responses or those of other elements, quite often the ganglion cell discharges.

One form of mass recording is the electroretinogram technique wherein one electrode is placed on the cornea and the other off the eye but near by. Electroretinograms contain the combined effects of diverse activities in the eye, and to make the records analytically useful requires skill and sophistication. At the hands of Granit and colleagues a remarkable amount of reliable information has been obtained.

The second form of mass recording is the recording of the optic nerve discharge outside the eye. This form, like the others, has its own advantages. By it, one can obtain a picture of the overall nature of the eye's output, which at times is a surer indication of what the various components in the retina contribute than the statistical inferences from microelectrode sampling of single elements. The fullest understanding demands the combined information from both mass and microelectrode techniques.

It is common when studying both periphery and center to simplify input so as to get at details of process. At the level of the optic pathway this sometimes means removal of the eye and application of electrical stimulation to the stump of the optic nerve. This method enables one to discover much about the behavior of the central portion of the optic pathway, quite impossible while including the complications produced by retinal activity. Once the character of cortical response to stimulation of the optic nerve (with the eye eliminated) is discovered, it can be compared with responses via stimulation of the retina and its attendant complications. This sort of comparison in the hands of Bishop and Bartley has provided a firm basis for some fundamental conclusions to be dealt with later.

2. *Impingement Variables*

Aside from intensity there are three main aspects of photic impingement fundamental to vision. They are retinal distribution (spatial factors), temporal factors, and spectral composition.

While in any visual situation, variables of space, time and composition may be manipulated in some combination, one can center on and manipulate a single class of variables to the virtual exclusion of the others. For our purposes we shall deal primarily with the effects of timing, and bypass much that is known about spatial interaction. It must be realized that the temporal distribution in the activity of the constituent elements of the pathway is a factor of prime importance. Magnitude of some aspect of neural activity is generally correlated with certain quantifiable aspects of sensory response, such as brightness. Magnitude in neural activity is generally expressed in such items as discharges per *unit time*. A given amount of energy produces ultimate effects depending upon how it is distributed in time. While this is illustrated, for example, in photochemistry, timing is also of great importance in neural activity. However, we must not suppose that timing in neural discharge has its only effect in determining the brightness aspect of sensory end result. Several other prominent aspects, such as hue, saturation and edge also depend upon timing.

Where the distribution in the temporal activities of neuron populations is involved, the concept of synchrony-asynchrony is relevant, and will be used in the presentation of findings.

It is only after one is able to depict some of the major temporal and intensive features of activity in the several segments of the pathway that he is ready to deal at length with spatial effects. Spatial interaction involves an elaboration of activity which is better avoided in the first stages of analysis.

C. THE RETINA

The first activity that must be taken into account in the retina is photochemical. This activity is the first discriminator in the response of the eye. With very brief photic impingements, duration as well as intensity is of consequence. The Bunsen-Roscoe law states that for very brief impingements, time and intensity are reciprocal. Below a critical duration, the impingement, as it is shortened, acts as though it were reduced in intensity. Thus there is a definite reduction in receptor activity with a consequent reduction in the sensory end result.

To understand how the over-all neural discharge of the retina is related to photic impingement, various investigators have studied eyes in which the effects of single receptors can be isolated. The question that follows is to what extent and in what way can this information be used. Careful work on single receptors has been in progress for a number of years. Early investigations into the manner in which a receptor responds to photic energy were made by Hartline (36, 37, 38, 39) and Hartline and Graham (40). At times it has proved tempting to use the nature of single receptor activity as a model for the over-all output of the retina, and as a determiner of some of the quantitative features of sensory end result. There seem to be several factors that preclude doing this.

First of all the thresholds of receptors differ widely. Impingements of low intensities activate a few receptors, and increasingly higher intensities activate progressively more, and over these ranges rates of discharge rise as intensity increases. Since, under most conditions, various receptors are activated at different points in their sensitivity ranges, the over-all effect is a complex combination of frequencies of discharge. As a part of this complexity all receptors are not equivalent in latency. This factor adds to asynchrony between the elements in action.

The greatest synchrony would occur in the response to a very brief isolated pulse or to the first pulse in a train. In other cases, this synchrony is lost and one cannot expect the discharge of the retina as a whole (or that part of the retina activated by a retinal image of restricted size) to mirror the activity of single receptors as recorded by microelectrodes.

The discharges of the ganglion cells into their axons are not all of the same type. Originally Hartline (36) discovered three general types: (*a*) an *on-off* response, occurring as a burst of impulses following the onset of a photic impingement and again at its termination; (*b*) the *maintained* response (another form of *on* response), a continuous discharge during the life of the impingement; and (*c*) the *off* response, a major volley of impulses following the termination of the impingement. In the Pecten, he (39) found only the *on* and *off* responses. In more recent literature, particularly on vetebrates, the terms *on, on-off* and *off* are more generally found. Some of these investigations have shown that the *on* and the *off* portions of the responses of the *on-off* elements may be dissociated. The discharge rate of the various types of neurons just mentioned to a train of photic pulses depends upon the rate of delivery of the pulses. As rate increases, the discharge rate (DR) increases up to a certain maximum and then descends (Enroth, 31; Dodt and Enroth, 29; Grüsser and Creutzfeldt, 32). This maximum thus occurs before CRP is reached. Some authors presume that brightness, the ultimate sensory end result, is determined by the rate of discharge in these retinal neurons.

Not all individual neurons of the same type reach their maximum response rate at the same point. This occurs over a wide range of photic pulse frequencies, let's say from 5 to 20 or more per second. (See Figure 2a.)

Some authors by averaging these various individual maxima have arrived at a theoretical mean maximum discharge to an input intermittency of about 10 per second. At first glance, this mean maximal DR might appear

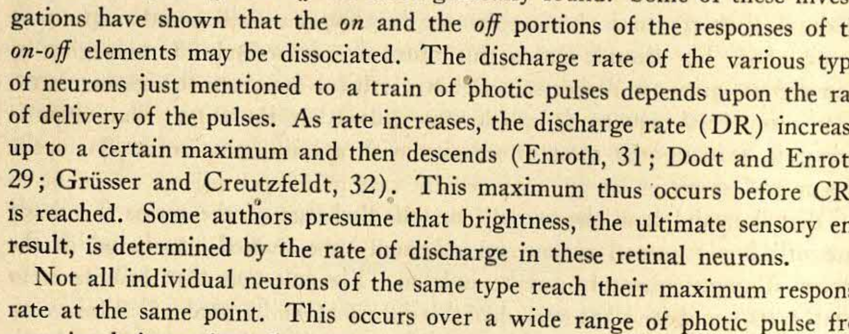


FIGURE 2

(a). Schematic representation of the relation between discharge rate (DR) and rate of photic intermittency for each of the three types of retinal neuron. A = *on* discharge; B = *off* discharge, and C = *on-off* discharge. Modified from Grüsser and Rabelo (33). (b). Schema showing the relation of latency, and pulse duration to intensity for *on* and *off* components of the optic nerve discharge. D_L = long pulse; D_S, short pulse. Left vertical dotted line = low intensity; right line, high intensity. (After Bartley and Bishop, 15.)

to be the basis for the sensory phenomenon, brightness enhancement[2] [Brücke-Bartley effect 5, 8, 32] which is often maximal at an impingement rate of 10 per second. That the retinal discharge is not the determiner of enhancement is based upon the following considerations: (*a*) there is no good synchrony in the timing of discharges to repeated photic impingements; and (*b*) it does not seem to be legitimate to use a mean maximum DR. Widely different maxima do not functionally average to a mean maximum as some have implied (33). For example, if one group of cells exhibits a mean maximal discharge to a photic pulse rate of 20 per second, another group to a rate of six per second, and a third to four per second, the averaging of these figures gives us 10 per second. Nevertheless one cannot suppose that a pulse rate of 10 per second can be substituted for the three separate pulse rates and produce the same effect as the three just mentioned. To speak of a mean maximum at an impingement rate of 10 per second would be to say that it made no difference whether the three separate rates or the single rate of 10 per second were used. Actually the over-all discharge in the present illustration could be expected to be a great deal less for a single input rate of 10 per second than would be required by the argument. These considerations preclude making direct inferences regarding over-all pathway responses from the action of any single cell or a few types of cells as a model. Only mass recording can make sure precisely what does happen.

1. *The Optic Nerve*

When the over-all discharge from the optic nerve is recorded, the main differentiations disclosed are the *on* and *off* components (Bartley and Bishop, 15). *On* responses are obtained with shorter and weaker impingements than *off* responses.

In a mid-range of both intensity and pulse duration, a second wave emerges in the *on* response. This was interpreted as evidence that stimulus conditions elicited responses from two receptor populations with differing latencies. This second wave is not found in the *off* response.

---

[2] The increased brightness of intermittent illumination over continuous illumination of the same intensity. Brightness enhancement produced by intermittent stimulation falls into two categories. The brightness enhancement usually studied is that produced by pulses repeated rapidly enough to produce a continuous visual field, of course with some degree of flicker. The second kind of brightness enhancement is produced when the photic pulse rate is low enough for the bright and dark phases of the visual field to be separate. In the first case, the observer gives brightness reports on a more or less mean value. In the second case, he reports on the bright phase as such. It must be recognized that these are two quite distinct situations and any investigator reporting on brightness values obtained when the pulse rate is considerably below 10 per second will be dealing with the second case and the values ought not to be unwittingly compared with those in case one as if they represented results from a true continuum.

Differences in the relation of response latency and photic intensity are manifested by the *on* and *off* responses. The latencies of *on* responses are generally shorter than those of the *off* responses. Whereas the *on* response latencies for long and short pulses converge to similar values at the high end of the intensity scale, the *off* response latencies for long and short pulses are similar for weak pulses and diverge in value as pulses are made more intense. Figure 2b shows this.

As pulses are repeated, the *on* and *off* responses change, and in different ways. The *off* response increases in amplitude as pulses are repeated. In contrast to this, *on* responses decline in size for the first few repetitions.

Recording from the optic nerve has not as yet been fully brought to bear on the problem at hand but it could be expected to tell us considerably more than we now know about the over-all pattern of the eye's output.

D. The Lateral Geniculate Body

The dorsal nucleus of the lateral geniculate body is made up of several layers, three in the cat, six in the monkey and man, and very little stratification in the rabbit. In the cat, layers A and B receive terminals of the tract fibers from the contralateral retina. The middle layer, A_1, is the terminus of fibers from the homolateral retina.

Not all fibers reaching the geniculate layers manifest the same conduction rates. The fastest groups reach layers A and A_1. Fibers from these relay to the striate cortex. The next slower fibers reach layer B and relay to the lateral nucleus of the thalamus. The next slower fibers go to the pretectal area and the very slowest go to the superior colliculus. But it is the fibers relaying to the cortex we are interested in.

In the lateral geniculate body, activity is still to be dealt with in terms of *on, on-off,* and *off* responses. All three forms of response are manifested. DeValois, Smith, Kitai and Karoly (28) report some elements act as *on* or *off* units when spectral stimulation is used. They act as *on* elements to one part of the spectrum, and as *off* elements to another part of the spectrum.

Not only are the two sets of responses identifiable, but response characteristics differ. Latencies of the *on* response 3-4 log units above threshold fall into three main groups, peaking at about 30, 140, and 380 msec. The *off* response peaks at the same intensities at 30 and 170 msec. Varying intensity of pulse dissociates the components of the *on-off* responses. That is, the two portions are affected in different ways. In most of the units, pulse intensity and response latency vary inversely for both *on* and *off* responses. In some cases, however, latency lengthens as a pulse becomes more intense and in some others no relation between the two is detectable.

Intense activation of the contralateral optic nerve by a brief electric shock produces two spikes (an *on* response), sometimes followed by a declining wave, recordable from an electrode just prior to the dorsal nucleus (22, 23). The second spike is conducted at about one-half the velocity of the first, although its threshold is about 2.5 times as high. Each of these spikes is followed by corresponding spikes beyond the dorsal nucleus synapse.

With inputs from common forms of photic stimulation one can expect an extended volley or volleys, with diversity in conduction rates of various elements producing them.

DeValois (26) reported in a discussion in a recent symposium that single cells in the cat's geniculate show the same pattern of behavior that Grüsser and Creutzfeldt (32) reported for retinal cells. That is, these cells fire repetitively in response to a single photic pulse. The first portion of the response is a high frequency burst of impulses which then tapers off in rate. If repetitive stimulation is employed, the impingement rate may be made rapid enough to include merely the burst. If the rate is made still greater, the succeeding impingements arrive soon enough to preclude the full bursts from occurring. This, in essence manipulates the total number of spikes per second in single elements here as in the retinal elements.

Under some conditions, maintained discharge is found in individual elements in the geniculate when examined by microelectrodes. That is, there is evidence that some cells are active in the absence of inputs from photic stimulation (DeValois, Jacobs and Jones, 27). Increments and decrements in illumination produce shifts in discharge rates of these cells at least for a while. This is not to be taken as evidence that the total or any large part of the total geniculate is active during absence of photic or other stimulation of the optic pathway. Over-all recording has not shown that it is. That this discharge is not likely autochthonous is reported by Arden and Söderberg (1).

Mass recording (Vastola, 53, 54) indicates that activation of the geniculate by a brief input from the optic nerve produces an after-positivity of 100-200 msec. and an inability of the cells to produce a new response.

After an initial phase of depression which becomes greatest at about 19 msec. following stimulation, recovery to 87 per cent of normal is reached at about 220 msec. (Bishop and Davis, 24). Full recovery requires about two seconds. This two stage recovery seems to largely account for the fact that maximal response at the cortex for a fully isolated stimulus is always greater than the maximal response to individual members in a train of stimuli. This difference has long been a puzzle since it represents a complication beyond the provisions inherent in any simple description of how the optic pathway works.

And finally the lateral geniculate produces potentiation at the cortex, through the summation of subliminal impulses arriving at the geniculate and lowering the threshold of the post-synaptic neurons (Chang, 25).

The substance of the foregoing findings and interpretations is simply that the lateral geniculate body plays a role in determining the rate at which any complete channel in the optic pathway may be reactivated. Although this rate is quite high during the very first few milli-seconds, depression or inhibition sets in and precludes immediate reactivation of the geniculate as a total. Inhibition is maintained for several seconds. The first stage of recovery would seem to participate in determining the periodicities in cortical response amplitude to be described in the next section. The second stage of recovery may account for the fact that temporally isolated stimuli evoke responses larger than the responses to members of any train of rapidly repeated stimuli even though they are timed for maximal effectiveness.

E. The Visual Cortex

The extended study of the optic pathway by stimulation of the optic nerve and recording from the visual area of the cerebral cortex was begun by Bishop and Bartley in the very early 1930's (2, 13, 14, 21). This antedated most of what has been discovered regarding certain other parts of the pathway. The information they obtained then has set the stage for the main issue dealt with in this paper, namely, the determination of the relation between the temporal features of impingement and sensory end result.

Bishop (21) initially attempted to see whether a detectable cortical response could be produced amid the ongoing and very irregular pattern of waves constituting the "spontaneous" activity of the cortex. He found response could be elicited but that it varied greatly in amplitude even when the strength of the repeated electrical impingement was kept constant. This led him to suspect the presence of a periodic variation in cortical receptivity to the optic nerve input. He checked this by a "tuning" procedure. First, the temporal interval between impingements was varied until a separation resulting in a constant amplitude of the cortical response was found. Secondly, the whole train of intermittent impingements was shifted back and forth in time, showing that some trains would produce uniformly small response, others uniformly medium-sized responses, and still others uniformly large responses.

Later the cortical effects produced by paired shocks to the stump of the optic nerve were studied by Bartley (3, 6). He found that the second member of the pair would not elicit a cortical response unless presented many milliseconds after the first, and would not elicit a full-sized response unless the separation was about 200 msec. (in the rabbit). This was taken as additional evidence for rhythmicity in the optic pathway.

Subsequent work has shown that very brief photic and electrical inputs produce essentially the same results. Trains of photic and electrical impingements produce cortical responses that are not greatly different in shape. The size of response elicited depends upon the phase relation between the impingement train and periodicity inherent in the visual pathway, despite the fact that the eye is involved in one case and not in the other. The same correspondence has been found to hold true when the effects of paired photic inputs hare compared to the effects of direct (electrical) stimulation of the optic nerve. These results strongly demonstrate that cortical response with its rhythmicity is not merely a quantitative replication of retinal action.

Maximal amplitudes of cortical response to intermittent photic stimulation (A_2)[3] in the monkey occur within a narrow frequency range (10.5 to 11.5 cycles per second) (34, 35). Halstead and colleagues also showed that intermittent impingements regularized the waves in the cortical record. Both of these effects are examples of producing synchrony in the active channel population, and were called "driving" by the authors. They also found that spectral impingements were more effective than whole-spectrum impingements in driving. This result is another example of increasing synchrony by reducing internal diversity in the impingement.

The cortex not only responds to the onset and continuance of a photic presentation with a specific potential, but to its termination (Bartley, 3). The visual cortex as well as the lower centers contain elements that can

---

[3] In various places, we allude to *maximal response*. These maxima are not all the same. There are at least three distinct maxima, depending upon the sets of conditions producing them. The cortical response to an isolated impingement is higher in amplitude than any cortical response under other conditions, hence we shall call its amplitude A_1. When repeated stimulation is used, a maximal response occurs which depends upon timing of the members in the impingement train. We shall call this A_2. A_3 occurs when rate of impingement intermittency exceeds the rate represented by the overall periodicity in the pathway. It is the amplitude of the cortical responses when repetitive conditions, in effect, approach the steady state, i.e., continuous illumination.

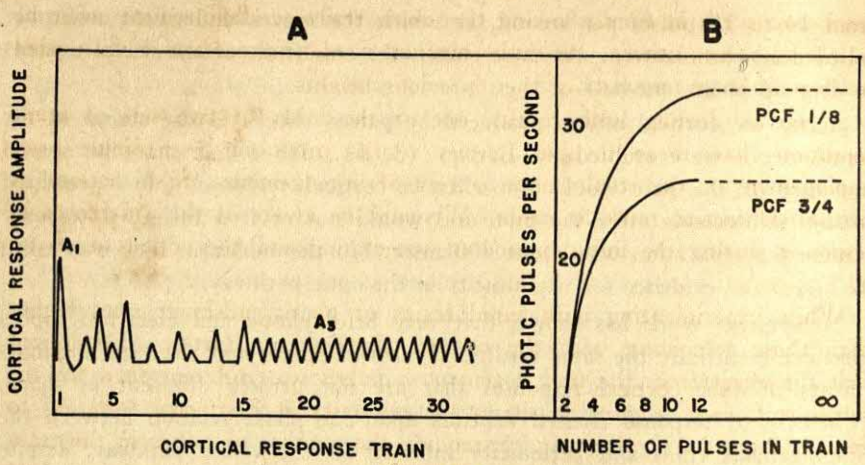


FIGURE 3

(a). Schema showing a typical response of the visual cortex (in the rabbit) to photic pulses of greater frequency than any single channel in the pathway can respond to. At first this inability shows up in various sized responses and complete failures to respond. Finally, after this reorganization period, all pulses are responded to equally (at amplitude A_3), no pulse being responded to by more than a fraction of the channels. (After Bartley, 3.) (b). Schematic curves to indicate the relation of photic pulse frequency required to just eliminate flicker, for trains of varying numbers of pulses from 2 to 12 to trains of indefinite length. Upper curve, for pulse-to-cycle fraction (PCF) 1/8; lower curve, PCF 3/4. (After Bartley, Nelson and Ranney, 17.)

be called *on* elements, and others which can be called *off* elements. In fact, microelectrode studies indicate five types of neurons (45, 46). The cortex can respond to the onset of a new pulse as early as 10 msec. following the termination of the preceding pulse, whereas two successive *on* responses can not be elicited without a much greater lapse of time. In fact, the second *on* response becomes as large as the first only when about 200 msec. have elapsed (in the rabbit). Thus quite separate channels for *on* and *off* effects are preserved all the way from the retina.

Results have been reported (Jasper, 43, 44) for human experiments that show cortical rhythmicity, and that all channels do not respond to all members of a train of photic pulses. Jasper recorded waves from the occipital cortex which tallied with photic pulse rates of 55 to 60 per second. He did not interpret this as evidence that the alpha rhythm of the cortex was driven beyond its typical rate of 8-13 per second since he discovered that the amplitude of the waves at 20 per second was about one-half of the amplitude at 10 per second and those at 40 per second one-fourth as high as at 10 per second. He showed that there were stages, for instance, at from 14 to 15 pulses per second, at which the waves underwent what he called desynchronization. At these stages, the amplitude of the waves tended to drop to about one-half of their previous heights.

So far in dealing with results, one or the other of two sets of input conditions have been involved, either results produced by a single brief impingement or the results once a train of impingements are in action. A further distinction must be made. We must be aware of the differences in responses during the initial period of stimulation and the period following it.

While experimenting with rapid trains of photic pulses at rates higher than those coinciding with the cortical rhythmicity, Bartley (3) noticed that the effect from the first dozen or so pulses was different than for the subsequent ones. Figure 3a shows the responses at first vary from A_1 to zero in a quite irregular way. Extending the number of pulses in the train eliminated the irregularity so that finally a uniform series of responses at amplitude A_3 was recorded. This initial period following the first pulse and during the delivery of the first few members of the train was interpreted as a time during which *reorganization* of responses among channels was achieved. When reorganization was complete, response waves were much smaller than A_1, but were all of the same size. The reorganization period may be described as a shift from the initial state in which all channels are activated simultaneously by the first pulse, and not all sufficiently recovered to respond to succeeding ones, to a state in which some channels respond to one pulse, and others to other pulses, all pulses finally eliciting action from equal numbers of channels. This amplitude is A_3.

F. RESUMÉ

There is not a uniform correspondence between repetitive rate in the delivery of photic pulses and discharge rate in single retinal units participating in the response. Under most conditions lack of synchrony between activity in the various units responding can be expected. Although certain patterns of individual unit activity can be isolated, the over-all retinal discharge is a composite in which many of the finer details of this activity are irrelevant for our purpose. We have had to discover what level of analysis or response is relevant in relation to the features of sensory end result to be accounted for.

At the lateral geniculate body, some of the main features of discharge in individual units found in the retina are preserved, but there a type of inhibition promotes periodicity in the activation of the channels of the optic pathway. Two forms of depression occur at least when a sizable number of parallel channels are involved. Most of this depression disappears in about 220 msec. in some species, and is all gone in about two seconds. The significance of this is not apparent till one determines the way the cortical response relates to impingement variables. Since we now know this, we can suppose that part of the periodicity in the receptivity of the channels in the pathway is determined at the geniculate by this depression. Recording from the geniculate with microelectrodes and disclosing simply what sample individual neurons do, seems to give no hint of this depression and thus no basis for attributing a role in channel periodicity to the lateral geniculate body.

At the cortex, there is no doubt about the existence of a periodicity in the ability of peripheral inputs to produce a response of a given size. The studies that have disclosed this are numerous and the demonstration has been repeated again and again.

We have not gone into detail in regard to the roles of the several types of cortical neuron. It would seem that the more definite roles of some of these show up best when spatial features of peripheral impingement are varied.

For our purposes, we have traced the temporal features of the results of both isolated and repetitive impingements through the pathway and have shown the major feature that has to be taken into account, namely the inherent periodicity and its role in the potential production of various temporal distributions in channel activity.

G. The Alternation of Response Theory

To organize and make sense out of the array of facts, Bartley (7, 11) formulated a statement consisting in an intertwining of information and inference which is called the alternation of response theory. This theory states that:

1. There is a fixed number of parallel channels in the optic pathway. In the optic nerve portion of the pathway, each fiber in the nerve is one segment of a channel.

2. After brief stimulation, there is a very short facilitation period, followed by a more extended period during which a channel cannot be reactivated, as evidenced by cortical response records. In the rabbit, this period is about 1/5 of a second. It is shorter in the cat, and it is about 1/10 of a second in man.

3. Depending on characteristics of impingement, the channel populations

may be activated in many different temporal distributions. A stimulus may activate all, many, or few channels at any one time, or the activation among channels may be rather uniformly distributed in time.

4. Brief intense impingements may activate all or nearly all channels virtually simultaneously. Channels set into action together tend to recover together, although recovery time may differ somewhat from channel to channel.

5. Activating all channels together produces a null period during which no response may be obtained from the pathway. With brief maximal impingements, the results are quite like those manifested by a single channel. By taking advantage of this, one may study the nature of a single channel.

6. The use of other patterns of stimuli activates the channels in various complex temporal relations to one another. Rates beyond the inherent rhythmicity of the pathway first produce a train of irregular-sized responses to a train of equal-sized impingements. Finally all channels are activated. When impingement is continuous instead of repetitive, soon as many channels go into activity as recover from activity at all instants. This is to say that uniform continuous impingeemnt results in uniform over-all response of pathway.

7. Over-all uniformity of response is affected by the principle of alternation of activity among the channels in the population. For instance, when a second impingement is delivered before all of the channels activated by the first impingement have recovered, the second response will depend upon the percentage of channels recovered. Only those channels that have recovered can be activated. The third impingement of the train will also be effective only in terms of the number of channels ready to respond. With an extended impingement sequence above the mean recovery rate of individual channels, some channels will respond to certain impingements, other channels to others. Very soon all impingements will find an equal number of channels ready to respond, and the net or over-all pathway activity will be uniform. Accordingly no continuous impingement, or repetitive train will ever be able to produce cortical responses with as great amplitudes as single isolated impingements.

8. Under many conditions, a correspondence between cortical-response amplitude and sensory end result (brightness) exists. The factors tending to produce amplitudinous responses in the cortex will produce high brightness. This is the connection between neurophysiology and sensation.

9. Several expectations should follow from the foregoing facts and interpretations. (*a*) Brief photic pulses repeated at their most effective rates should

produce higher brightnesses than longer ones (18, 19). (*b*) Weak photic pulses should not produce brightness enhancement (9). (*c*) Lower PCFs (pulse-to-cycle fractions) should be more effective than higher ones when maximally effective repetitive rates are used (18, 19). (*d*) No continuous illumination can produce as high brightness as certain brief trains of repeated pulses or isolated pulses. (*e*) Brightness enhancement may be absent when the visual target image covers only a very small fraction of the retina. Too many other receptors (thus parallel and collateral channels) are left free to be stimulated sporadically or not at all and so to interfere with the degree of synchrony required of active elements in the restricted region of the retina. That is, the timing of the impingement may have too little effect upon the optic pathway as a whole. (*f*) When few photic pulses constitute a train, greater separations between pulses will be required to produce flicker than when the train is of unlimited duration. Since observers cannot make response distinctions between the first part of a train and the train as a total they must be presented with trains of different numbers of pulses (17). (See Figure 3b.) (*g*) Brightness enhancement is reduced by the existence of stray illumination in the eye (20). (*h*) Separating photic pulses in a train by longer and longer null periods ought to increase brightness until it becomes equivalent to the brightness produced by a single pulse (50). (*i*) Since it has been repeatedly observed in brightness enhancement that the color characteristics of the target begin to change as photic illumination is made intermittent, the alternation of response concept seems to apply to color as well as brightness. This lead has been followed in several investigations in our laboratory with positive results (16, 48, 49). (*j*) Various other predictions can be made that will not be enumerated here.

H. SIMILAR THEORIES AND CONCEPTS IN NON-VISUAL SYSTEMS

Alternation of response is a feature of the complex optic pathway mechanism which accounts for brightness enhancement and various other subtle and paradoxical features of sensory response. It seems that the principle applies not only in the visual mechanisms but underlies certain phenomena in other sense modalities. These are phenomena which have been independently studied in audition, kinesthesis and touch, and give support to the plausibility of the principle operating in the visual modality.

In audition, certain complexities have arisen in relating vibration frequencies of the stimulus, the maximum rate at which single nerve fibers can discharge, the wave frequencies seen in records of the auditory nerve, and pitch as a sensory end result. Two factors, the "place" in the organ of

Corti and the "frequency" of responses are involved. To account for known ability to discriminate between differences in frequency even when input frequency is high (in the thousands) requires either that single nerve fibers respond one-to-one to high frequencies or various fibers redistribute their action in relation to one another. This latter possibility was suggested by Wever and Bray (56) and Wever (55) and has become known as the *volley theory*.

The visual activity that we have described and labelled the alternation of response theory seems to be an analogue of this in certain important respects. One main difference, of course, is the fact that the volley theory is a peripheral response theory, and the alternation of response theory refers to central nervous function.

There is evidence from several other sources for the general rotation or alternation principle in the nervous system. Echlin and Fessard (30) have described activity of the kinesthetic system which is relevant. The highest discharge frequency they obtained from a single "stretch" receptor in the frog was 330 per second at 22°C. This compares favorably with Matthews' (47) values of 290 per second at 16°C. However, in recording from the muscle of frog, frequencies as high as 440 per second have been recorded at 17°C. Cooling the preparation with iced Ringer solution for 30 seconds did not eliminate the 440 per second discharge up to about the point at which the discharge failed altogether. This means that the high of discharge was a composite of single elements not in phase and not each responding at the high rate seen in the muscle record. If the rate were a direct expression of receptor discharge frequencies, one would have expected cooling to have reduced the frequency recorded in the records.

Rose and Mountcastle (51) describe a relation between the capacity of the tactile system as a whole and the capacity of a single element to respond which seems to be similar in principle to the alternation of response in the visual system. They point out that the results obtained from studying responses to paired stimuli do not reflect the capacity of the system as a whole to respond to trains of stimuli. At certain frequencies, the neuron follows the stimulus rate with a one-to-one relation. When trains of stimuli are presented at higher rates, the individual neurons still respond at their original or apparently intrinsic rates. Response to every member of the impingement train is to be found because some neurons respond to some members of the train and others to others. While this provides for response no matter how high the stimulus rate, it also means that the intensity of over-all response per unit of time is less than the intensity when all the available neural elements respond together in synchrony, as to a single brief impingement. This is of course very like the description of the visual system given by alternation of response theory.

Rose and Mountcastle also point out that whereas the first impingement in a train evokes a response, immediately subsequent members of the train are irregularly responded to. That is, for a period some impingements do not elicit any response while others do. If the train is sufficiently prolonged all the stimuli are finally responded to. This description too, is quite like the description of the reorganization period occurring in the visual pathway.

Still another similarity between the tactile or somatic system and the visual system is evident. The discharge from receptors, and from the subsequent neurons in the pathway are short trains themselves when single stimuli are used, but these trains each reduce to a single impulse when stimuli are repeated in trains at high rates. This reduction occurs in the tactile system when rate reaches a neighborhood of 10 to 15 per second, under the circumstances used in Mountcastle and Rose's experiments (51). This is reminiscent of the conditions producing CRP in the visual system.

I. Summary and Conclusions

It is customary to move step by step from the periphery to the central nervous system in tracing what happens in the optic pathway. This procedure is not always the most appropriate, if we want to know as soon as possible something about what takes place in the neural structures that immediately underly sensory end results. One would expect to be able to make connections between peripheral and central events without first exploring all the intermediate activities in the pathway. This would disclose a set of broad relations between the two. Although such an analysis would not deal with fine detail it would provide the framework upon which the finer detail obtained later could be placed. Bishop and Bartley's original work during the early thirties provided the beginning of such a framework. Broad temporal relationships between the peripheral input and cortical response possibilities were worked out. Though the findings lacked a great deal of fine detail, they did yield knowledge of a general pattern of pathway activity that seems to be well confirmed.

Since the work in the thirties, considerable improvement and increased use of the microelectrode technique has taken place. Since it is now common to record from single neurons, it has become quite customary to think in terms of the fine detail that this technique provides. Nevertheless it is thinkable that use of this technique alone, simply moving single-unit analysis from periphery to center, would not yet have disclosed the broad over-all patterning of activity we now know to exist in the pathway. Since we now possess a broad outline regarding the relation of the receptivity of the cortex to peripheral inputs, we should now be better able to evaluate and utilize microelectrode findings.

The present paper has indicated a number of facts about temporal features of response in the retina, optic nerve, lateral geniculate body, and in the cortex which seem to support the theory that sensory end result produced by photic input to the eye depends upon various features of timing. A theory based upon these facts and certain interpretations, called the alternation of response theory was restated. This provides an understanding of the neurophysiological basis for brightness, particularly brightness enhancement, and of several other features of vision. A number of predictions which stem from the theory in regard to brightness and brightness enhancement have been confirmed. In fact, in no instance in which any test has been made has negation resulted.

The final section of the paper pointed out that individual channels in the pathways serving audition, touch, and kinesthesis manifest a form of alternation of response akin to that theorized for the optic pathway.

REFERENCES

1. ARDEN, G. B., & SÖDERBERG, U. The transfer of optic information through the lateral geniculate body of the rabbit. *In* Rosenblith, W. A. (*Ed.*), *Sensory Communication.* Cambridge, Mass.: Technology Press, 1961.
2. BARTLEY, S. H. Action potentials of the optic cortex under the influence of strychnine. *Amer. J. Physiol.,* 1933, **103**, 203-212.
3. ———. Temporal and spatial summation of extrinsic impulses with intrinsic activity of the cortex. *J. Cell. & Comp. Physiol.,* 1936, **8**, 41-62.
4. ———. Subjective flicker rate with reference to critical flicker frequency. *J. Exp. Psychol.,* 1938, **22**, 388-394.
5. ———. Some factors in brightness discrimination. *Psychol. Rev.,* 1939, **46**, 337-358.
6. ———. The relation between cortical response to visual stimulation and changes in the alpha rhythm. *J. Exp. Psychol.,* 1940, **27**, 624-639.
7. ———. Visual sensation and its dependence upon the neurophysiology of the optic pathway. *Biol. Symposia,* 1942, **8**, 87-106.
8. ———. Brightness enhancement in relation to target intensity. *J. of Psychol.,* 1951, **32**, 57-62.
9. ———. Intermittent photic stimulation at marginal intensity levels. *J. of Psychol.,* 1951, **32**, 217-223.
10. ———. Principles of Perception. New York: Harper, 1958.
11. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *AMA Arch. Ophthal.,* Pt. II, 1958, **60**, 775-791.
12. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. *In* Jung, R. (*Ed.*), *The Visual System:*

Neurophysiology and Psychophysics Symposium, Freiburg. Heidelberg: Springer-Verlag, 1961.

13. BARTLEY, S. H., & BISHOP, G. H. The cortical response to stimulation of the optic nerve of the rabbit. *Amer. J. Physiol.,* 1933, **103**, 159-172.
14. ———. Factors determining the form of the electrical response from the optic cortex of the rabbit. *Amer. J. Physiol.,* 1933, **103**, 173-184.
15. ———. Some features of the optic nerve discharge in the rabbit and cat. *J. Cell. & Comp. Physiol.,* 1942, **19**, 78-93.
16. BARTLEY, S. H., & NELSON, T. M. Certain chromatic and brightness changes associated with rate of intermittency of photic stimulation. *J. of Psychol.,* 1960, **50**, 323-332.
17. BARTLEY, S. H., NELSON, T. M., & RANNEY, J. E. The sensory parallel of the reorganization period in the cortical response in intermittent retinal stimulation. *J. of Psychol.,* 1961, **52**, 137-147.
18. BARTLEY, S. H., PACZEWITZ, G., & VALSI, E. Brightness enhancement and the stimulus cycle. *J. of Psychol.,* 1957, **43**, 187-192.
19. BARTLEY, S. H., VALSI, E., & BOURASSA, C. Further manipulation of brightness enhancement. *J. of Psychol.,* 1959, **48**, 47-55.
20. BARTLEY, S. H., & WILKINSON, F. R. Brightness enhancement when entoptic stray light is held constant. *J. of Psychol.,* 1952, **33**, 301-305.
21. BISHOP, G. H. Cyclic changes in excitability of the optic pathway of the rabbit. *Amer. J. Physiol.,* 1933, **103**, 213-224.
22. BISHOP, G. H., & O'LEARY, J. L. Electrical activity of the lateral geniculate of cats following optic nerve stimuli. *J. Neurophysiol.,* 1940, **3**, 308-322.
23. ———. Factors determining the form of the potential record in the vicinity of synapses of the lateral geniculate body. *J. Cell. & Comp. Physiol.,* 1942, **19**, 315-331.
24. BISHOP, P. O., & DAVIS, R. The recovery of responsiveness of the sensory synapse in the lateral geniculate. *J. Physiol.,* 1960, **150**, 214-238.
25. CHANG, H. J. Cortical response to stimulation of lateral geniculate body and the potentiation thereof by continuous illumination of the retina. *J. Neurophysiol.,* 1952, **15**, 5-26.
26. DEVALOIS, R. L. His discussion found in Reference 11.
27. DEVALOIS, R. L., JACOBS, G. H., & JONES, A. E. Effects of increments and decrements of light on neural discharge rates. *Science,* 1962, **136**, 986-987.
28. DEVALOIS, R. L., SMITH, C. J., KITAI, S. T., & KAROLY, A. J. Response of single cells in monkey lateral geniculate nucleus to monochromatic light. *Science,* 1958, **127**, 238-239.
29. DODT, E., & ENROTH, C. Retinal flicker response in cats. *Act. Physiol. Scand.,* 1954, **30**, 375-390.
30. ECHLIN, F., & FESSARD, A. Synchronized impulse discharges from receptors in the deep tissues in response to a vibrating stimulus. *J. Physiol.,* 1938, **93**, 312-334.
31. ENROTH, C. Spike frequency and flicker fusion frequency in retinal ganglion cells. *Act. Physiol. Scand.,* 1953, **29**, 19-21.
32. GRÜSSER, O.-J., & CREUTZFELDT, O. Eine neurophysiologische Grundlage des Brücke-Bartley-Effektes: Maxima der Impulsfrequenz retinaler und corticaler Neurone bei Flimmerlicht mittlerer Frequenzen. *Pflüg. Arch.,* 1957, **263**, 668-681.
33. GRÜSSER, O.-J., & RABELO, C. Reaktionen retinaler Neurone nach Lichtblitzen I. *Pflüg. Arch.,* 1958, **265**, 501-525.
34. HALSTEAD, W. C., KNOX, G. W., & WALKER, A. E. Modification of cortical activity by means of intermittent photic stimulation in the monkey. *J. Neurophysiol.,* 1942, **5**, 349-355.

35. HALSTEAD, W. C., KNOX, G. W., WOOLF, J. I., & WALKER, A. E. Effects of intensity and wave length on driving cortical activity in monkeys. *J. Neurophysiol.*, 1942, **5**, 483-486.

36. HARTLINE, H. K. A quantitative and descriptive study of the electric response to illumination of the arthropod eye. *Amer. J. Physiol.*, 1928, **83**, 466-483.

37. ———. Intensity and duration in the excitation of single photoreceptor units. *J. Cell. & Comp. Physiol.*, 1934-34, **5**, 229-247.

38. ———. Response of single optic nerve fibers of the vertebrate eye to illumination of the retina. *Amer. J. Physiol.*, 1938, **121**, 400-415.

39. ———. Discharge of impulses in the optic nerve of Pecten in response to illumination of the eye. *J. Cell. & Comp. Physiol.*, 1938, **11**, 465-478.

40. HARTLINE, H. K., & GRAHAM, C. H. Nerve impulses from single receptors in the eye. *J. Cell. & Comp. Physiol.*, 1932, **1**, 277-295.

41. HECHT, S. The photochemical nature of the photosensory process. *J. Gen. Physiol.*, 1919-20, **2**, 229-246.

42. ———. The nature of the photoreceptor process. *In* Murchison, C. (*Ed.*), *Handbook of General Experimental Psychology*. Worcester: Clark Univ. Press, 1934.

43. JASPER, H. H. Electrical signs of cortical activity. *Psychol. Bull.*, 1937, **34**, 441-448.

44. ———. Reports and communications. *XI Cong. Internal. Psychol.*, 1937, 226.

45. JUNG, R. Neuronal integration in the visual cortex and its significance for visual information. *In* Rosenblith, W. A. (*Ed.*), *Sensory Communication*. Cambridge, Mass.: Technol. Press, 1961.

46. JUNG, R., v. BAUMGARTEN, R., & BAUMGARTEN, G. Mikroableitung von einzelnen Nervenzellen im optische Cortex der Katze: die licht-activativerten B-neurone. *Arch. Psychiat. Nervenkrankh.*, 1952, **189**, 521-539.

47. MATTHEWS, B. H. C. The response of a single end organ. *J. Physiol.*, 1931, **71**, 64-110.

48. NELSON, T. M., & BARTLEY, S. H. The role of PCF in temporal manipulations of color. *J. of Psychol.*, 1961, **52**, 457-477.

49. NELSON, T. M., BARTLEY, S. H., & MACKAVEY, W. R. Responses to certain pseudoisochromatic charts viewed in intermittent illumination. *Percept. & Mot. Skills*, 1961, **13**, 227-231.

50. NELSON, T. M., & BARTLEY, S. H. The relation between brightness produced by repeated photic pulses and the length of the null period between them. (To be published.)

51. ROSE, J. E., & MOUNTCASTLE, V. B. Touch and kinesthesis. *In* Magoun, H. W. (*Ed.*), *Handbook of Physiology, Sec. 1, Vol. 1.* Washington: Amer. Physiol. Society, 1959.

52. SÉGAL, J. Le méchanism de la vision en lumière intermittent. *J. Psychol. Norm. & Path.*, 1939, **36**, 451-539.

53. VASTOLA, E. F. Steady potential responses in the lateral geniculate body. *EEG & Clin. Neurophysiol.*, 1955, **7**, 557-567.

54. VASTOLA, E. F. After-positivity in lateral geniculate body. *J. Neurophysiol.*, 1959, **22**, 258-272.

55. WEVER, E. G. Theory of Hearing. New York: Wiley, 1949.

56. WEVER, E. G., & BRAY, C. W. Present possibilities for auditory theory. *Psychol. Rev.*, 1930, **37**, 365-380.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 145-151.

FAMILIARIZATION WITH THE TEST APPARATUS AS A FACTOR IN THE REINFORCING EFFECT OF CHANGE IN ILLUMINATION*[1]

Yale University and University of Connecticut

R. N. LEATON, D. SYMMES, AND H. BARRY, III

A. INTRODUCTION

The apparent reinforcing effect of the onset of dim illumination has been repeatedly demonstrated in the rat (see 4 for a review of these studies). Some investigators have suggested that stimulus change or novelty is the reinforcing agent (9, 11), others that the opportunity for visual scanning of the environment contributes to the reinforcement (12, 14), while still others compare the phenomenon to organic-drive-related reinforcement (6).

The degree of stimulus change or novelty represented by light onset should be greater if *S*s have previously become familiar with other aspects of the test environment than if the entire situation is equally novel; thus the reinforcing effect should be greater if it is related to stimulus change, per se. On the other hand, prior familiarization should decrease the reinforcement effect if visual scanning is the major agent of reinforcement. Familiarization should have little effect if the reinforcement is comparable to organic-drive-related reinforcements.

Results by Girdner (7) and Appel and Hurwitz (1) have suggested that the reinforcing effect of light onset may be enhanced by pretest familiarization with the test apparatus. Since changes in the animals' level of activity were not adequately controlled in these studies, which tested rate of pressing a single lever which turned on a light, it is difficult to evaluate the contribution of the familiarization to the reinforcement apart from any effect on the unconditioned rate of responding.

The offset of dim illumination constitutes a stimulus change while decreasing the opportunity for visual scanning of the environment. Therefore

---

* Received in the Editorial Office on August 29, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported, in part, by Neal E. Miller's Research Grant MY-2949 from the National Institute of Mental Health, United States Public Health Service, and was carried out during the senior author's tenure of a Predoctoral Fellowship from the National Institute of Mental Health. Results were reported at the annual meeting of the American Psychological Association, 1961.

studies which have shown a reinforcing effect of light offset, in the same situation where a reinforcing effect of light onset was found (3, 13), support the theory that stimulus change is the common reinforcing factor in both. On the other hand, failures to find a reinforcing effect of light offset (2, 8, 14) are more consistent with the theory that the reinforcing effect of light onset is due to the opportunity for visual scanning.

The present experiment was designed to study the relationship between the degree of familiarity with the test apparatus prior to light-change contingency and the magnitude and course of the reinforcing effects of both light onset and light offset. A two-lever apparatus was used in order to evaluate the reinforcing effects apart from any differences in the unconditioned rate of responding.

B. METHOD

1. *Subjects*

The *S*s were 36 naive, male albino rats, of Sprague-Dawley strain, approximately 100 days old at the beginning of the experiment. They were housed in individual cages with food and water available at all times.

2. *Apparatus*

The test box was 12 inches long, 9½ inches wide, and eight inches high, constructed of plywood walls painted flat black, a hardware cloth floor and a white Plexiglas ceiling. The only illumination was provided by a 7½-watt bulb, mounted 12 inches above the translucent ceiling, which produced 0.27 foot-candles of illumination at the floor of the box. Centered on each end wall of the box was a flat aluminum lever two inches above the floor, four inches wide and protruding 1½ inches from the wall. The discriminability of the levers was enhanced by perforations in one of them. A ½ inch depression of either lever, requiring a force of from 10-15 gm., operated a noiseless mercury switch connected to remotely located electronic counters and timers. The light in the test box could be turned on or off by either lever for as long as the lever remained depressed. The test box was housed in a light-tight, ventilated chamber, which provided effective insulation from noise.

3. *Procedure*

After several days of habituation to their living cages, *S*s were handled by *E* for several minutes on each of two successive days. At this time, by random selection, half were assigned to the light-onset condition and half to the light-offset condition. Both groups were further randomly divided into three subgroups given different amounts of familiarization to the test box (0, two or seven days) prior to the introduction of the light-change contingency. Six experimental groups of six *S*s each were thus formed. On each of the next seven days the *S*s in the seven-day groups were given a 20-minute familiarization session in the test box during which the light remained continuously off (for the light onset group) or continuously on (for the light offset group). The *S*s assigned to the two-day groups were placed in the box only on the last two of these days; the *S*s given no prior familiarization were not placed in the box at all during these seven days. On each of these days *S*s not tested were given the same handling by *E* as was involved in testing the other *S*s.

Following these seven days all *S*s were given a 20-minute test session on each of 12 consecutive days during which pressing one lever turned the light on (if the box was normally dark) or off (if the box was normally lighted) for as long as the lever remained depressed. The other lever continued to have no effect. For each *S* the positive and neutral levers were selected by chance. In order to equate for any possible differences between the levers, each served as the positive lever for half the *S*s and as the neutral lever for half within each experimental group.

C. RESULTS

Figure 1 shows the mean number of responses per session on the positive and neutral levers for the three groups in the light-onset condition. An analysis of variance of these data showed a highly reliable difference between the positive and neutral levers ($F = 70.43$; $df = 1, 15$; $p < .001$), providing further evidence for the reinforcing effect of the onset of dim illumination. The difference between the positive and neutral levers (preference) declined for the first few sessions for the two groups given prior familiarization with the test box. On the other hand, the zero-familiarization group showed only a slight preference for the positive lever during the first test session, but this preference increased progressively for the first few sessions. This effect of familiarization was revealed statistically by a significant groups × sessions interaction with preference ($F = 4.01$; $df = 22, 165$; $p < .001$). Thus the groups differed significantly with respect to magnitude of positive-lever preference on successive sessions. Since all three groups showed a similar preference for the last seven sessions, there was no reliable over-all difference in preference among the groups for the 12 test sessions combined ($F = 1.28$; $df = 2, 15$).

The onset groups did differ reliably in total lever pressing ($F = 7.88$; df

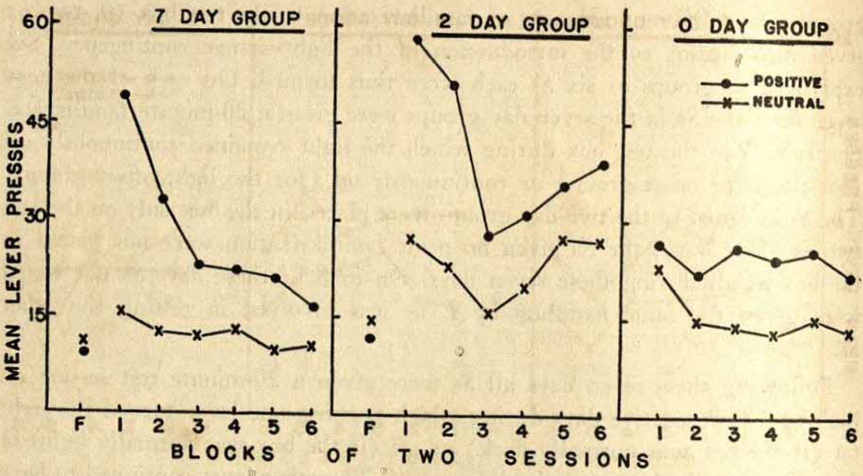


FIGURE 1
MEAN LEVER PRESSES PER SESSION FOR LIGHT ONSET. "F" INDICATES LAST FAMILIARIZATION SESSION

$= 2, 15; p < .005$). This difference is attributable to the 2-day group's higher rate of pressing both levers, since the seven-day and 0-day groups had almost identical over-all lever pressing rates. Thus while the magnitude and course of the preference for the light-onset lever were very similar for the two-day and seven-day groups, the additional days of familiarization were reflected in the unconditioned rate of responding, and the reinforcing effect for the two-day group would have been overestimated had a one-lever box been used.

Figure 2 shows the corresponding data for the light-offset condition. The difference between the positive and neutral levers was significant at $p < .001$ ($F = 21.51$; $df = 1, 15$). Thus a reliable reinforcing effect of the offset of dim illumination was demonstrated. As can be seen from the figure, a consistent preference for the positive lever was found only in the two groups given prior familiarization with the test box. The three groups differed significantly with respect to positive-lever preference ($F = 3.60$; $df = 2, 15$; $p = .054$). Thus the effect of familiarization with the test box for the light-offset condition was to reveal a preference for the positive lever which did not develop without such familiarization.

D. DISCUSSION

The present study shows that the reinforcing effects of light onset and light offset are, in part, a function of prior familiarization with the test

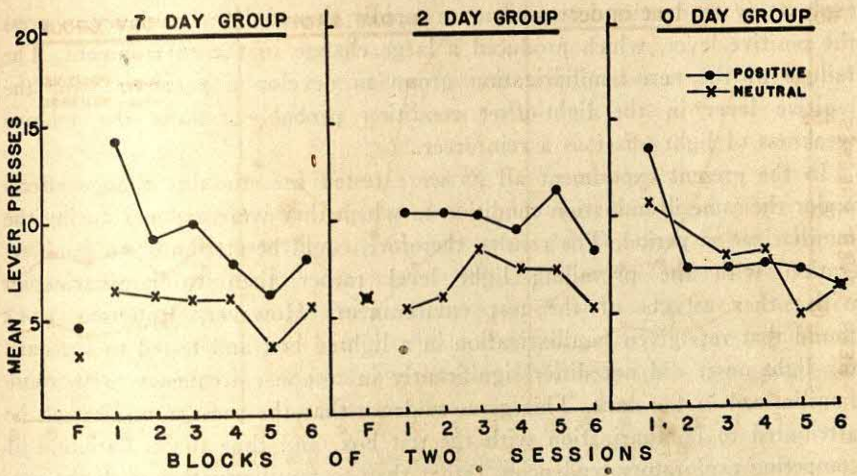


FIGURE 2

Mean Lever Presses Per Session for Light Offset. "F" Indicates Last Familiarization Session

environment. These results support a novelty interpretation of the reinforcing effects of changes in illumination. A decline in novelty accounts for the typical decrease in lever-pressing for stimulus change over a series of test days. For an animal already familiar with the test box, the light onset or offset will be the most novel feature of the situation; thus the reinforcing effect of light change should be maximal on the first day it is introduced and thereafter gradually decline. This result was found in the present experiment for the light-onset groups given two or seven prior familiarization days. Although two days of familiarization were sufficient for a full development of the preference for the light-onset lever, the seven-day group's lower rate of pressing both levers would seem to indicate that the additional five familiarization days were effective in reducing general activity.

For groups given no familiarization to the test box prior to light-change contingency, the entire situation is equally novel. A variety of other exploratory tendencies, including pressing the neutral lever, will compete with the tendency to press the lever which produces light-change. This hypothesis could account for the finding that the groups with no prior familiarization failed to show a positive-lever preference on the first test session. The zero-familiarization group in the light-onset condition showed a progressive increase in positive-lever preference for the first few test days due to a decrease in rate of pressing the neutral lever. This seems to indicate that competing exploratory tendencies decreased more rapidly than did the tendency to press the positive lever, which produced a large change in the environment. The failure of the zero-familiarization group to develop a preference for the positive lever in the light-offset condition probably reflects the relative weakness of light offset as a reinforcer.

In the present experiment all *S*s were tested for stimulus change effects under the same illumination condition to which they were exposed during the familiarization period. The results, therefore, could be attributed to familiarization with the prevailing light level rather than to familiarization with other aspects of the test environment. However, Robinson (14) found that rats given familiarization in a lighted box and tested in the dark for light onset did not differ significantly in response frequency from those familiarized in the dark. This gives evidence that the present results can be attributed to familiarization with the test box, and thus to the satiation of competing exploratory tendencies, rather than to familiarization with the prevailing light condition.

A number of studies have introduced the light-onset contingency for all animals on the first session in the test box, thus allowing for no prior familiarization with the apparatus (e.g., 6, 10). Such studies have typically found a progressively increasing rate of pressing for light onset. On the other hand, experiments which have used prior familiarization with the test box (e.g., 7, 11) have shown, with only one known exception (5), a progressively decreasing rate both between and within test sessions. The present results suggest that the difference in these results is due to the presence or absence of prior familiarization.

The present experiment clearly shows that the offset of dim illumination, as well as the onset, can reinforce lever pressing. The demonstration of a reinforcing effect of light offset makes interpretation in terms of stimulus change more convincing. The available evidence in the experimental literature has indicated that such an effect is difficult to demonstrate. The offset effect in the present experiment was much smaller than the onset effect; thus the small relative magnitude of the reinforcing effect of light offset, as well as the fact that the effect appeared only in the groups given prior familiarization, could partially explain the lack of consistent results in the literature. One of the possible reasons for the large difference in the light-onset and light-offset preference could be the much lower level of total lever-pressing in the light-offset condition. This suggests that the lighted box tended to elicit visual exploration, crouching, or some other activity incompatible with pressing either lever.

E. SUMMARY

The effect of contingent dim light onset and offset on lever pressing by albino rats was studied in groups with 0, two and seven days of prior familiarization with the test box. Reinforcing effects of the stimulus change were inferred from preference for a "positive," illumination-change lever in comparison with a "neutral" lever. Both light onset and light offset were found to be reinforcing for rats familiar with the test box, but the offset effect was much smaller than the onset effect. Rats without prior familiarization showed no reinforcing effect of light onset or offset during the first few test sessions, but the light onset group developed a significant preference for the positive lever after three to five sessions. The light offset group showed no such delayed reinforcing effect.

REFERENCES

1. APPEL, J. B., & HURWITZ, H. M. B. Studies in light reinforced behavior: IV. Effects of apparatus familiarization. *Psychol. Rep.*, 1959, **5**, 355-356.
2. BARNES, G. W., & KISH, G. B. Behavioral effects of the cessation of weak light energy. *Amer. Psychologist*, 1957, **12**, 411. (Abstract).
3. BARRY, H., III, & SYMMES, D. Reinforcing effects of illumination change in different phases of the rat's diurnal cycle. *J. Comp. & Physiol. Psychol.*, in press.
4. BERLYNE, D. E. Conflict, Arousal, and Curiosity. New York: McGraw-Hill, 1960.
5. FORGAYS, D. G., & LEVIN, H. Discrimination and reversal learning as a function of change of sensory stimulation. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 191-197.
6. ———. Learning as a function of change of sensory stimulation: Distributed vs. massed trials. *J. Comp. & Physiol. Psychol.*, 1961, **54**, 59-62.
7. GIRDNER, J. B. An experimental analysis of the behavioral effects of a perceptual consequence unrelated to organic drive states. Unpublished doctoral dissertation, Duke Univ., 1953.
8. HURWITZ, H. M. B. Conditioned responses in rats reinforced by light. *Brit. J. Anim. Behav.*, 1956, **4**, 31-33.
9. KISH, G. B. Learning when the onset of illumination is used as reinforcing stimulus. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 261-264.
10. LEVIN, H., & FORGAYS, D. G. Learning as a function of sensory stimulation of various intensities. *J. Comp. & Physiol. Psychol.*, 1959, **52**, 195-201.
11. MARX, M. H., HENDERSON, R. L., & ROBERTS, C. L. Positive reinforcement of the bar-pressing response by a light stimulus following dark operant pretests with no after-effect. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 73-76.
12. MOWRER, O. H. Learning Theory and Behavior. New York: Wiley, 1960.
13. ROBERTS, C. L., MARX, M. H., & COLLIER, G. Light onset and light offset as reinforcers for the albino rat. *J. Comp. & Physiol. Psychol.*, 1958, **51**, 575-579.
14. ROBINSON, J. S. Light onset and termination as reinforcers for rats living under normal light conditions. *Psychol. Rep.*, 1959, **5**, 793-796.

Department of Psychology
Yale University
New Haven, Connecticut

Published as a separate and in *The Journal of Psycholgoy,* 1963, **55**, 153-163.

VISUAL ACUITY UNDER CONDITIONS OF INTERMITTENT ILLUMINATION PRODUCTIVE OF PARADOXICAL BRIGHTNESS*[1]

Department of Psychology, Michigan State University

S. HOWARD BARTLEY, THOMAS M. NELSON, AND ELAINE M. SOULES

A. INTRODUCTION

The ability to read small type, or deal with fine spatial details of objects is based upon what is called visual acuity. Visual acuity is defined as the ability to resolve various portions of a target when minimal angular separation between the portions is a limiting factor. Measurement simply involves ascertainment of the minimal perceivable separation of two target components which is expressed as the reciprocal of visual angle.

Research into visual acuity has traditionally involved two types of physical manipulation in a static visual field. One procedure measures changes effected by various spatial distributions of intensity; the other measures changes effected by variations in intensity per se, i.e., the ratio of transmittance or reflectance to non-transmittance or non-reflectance.

It is not surprising, because of this, that acuity is conceptualized largely in terms of intensive distribution of luminance in the target. Visual acuity might well be defined in terms of the distribution of luminance on the retina, for this distribution differs from the target specifications and at the same time more truly represents the pattern of energy given the organism. It is true, however, that biological theorists do tend to account for acuity in terms of the fineness of the retinal mosaic.

There are limitations to any theory taking account only of variables in a static field. This became obvious when visual acuity for moving targets was measured. It was discovered that dynamic visual acuity, acuity under conditions of target displacement, is not predictable from visual acuity in static situations.

Study of the role of temporal factors in visual acuity has also enlarged knowledge of acuity. For example, it has been discovered that acuity becomes poorer as exposure time is shortened. Little or no significance has been at-

---

* Received in the Editorial Office on August 29, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

tached to this fact however, perhaps either because the findings have to do with conditions lying outside the "practical" range or because they are too much in line with other findings involving shortened exposure to merit special theoretical attention. In general, shortened exposure reduces or impoverishes any sensory discrimination.

A recent resurgence of interest in the effect of temporal variables upon visual acuity led to the study of visual acuity in intermittently illuminated fields. Senders (13), Nachmais (11), and Gerathwohl and Taylor (4) have all obtained results they did not expect. Senders showed that field brightness induced by intermittent stimulation does not co-vary with visual acuity in any simple way. Gerathwohl and Taylor have shown visual acuity to be poorer under intermittent stimulus conditions. They were not able to tie the effect to particular features of intermittency however. Nachmais related acuity changes to brightness changes induced by intermittent stimulation probably at rates above CFF.

The present study aims to discover whether brightness increments induced by intermittent stimulation have the same effect upon visual acuity as brightness changes induced by increments in intensity. Unlike the others, our aim will be to relate the findings to more general neural processes of the visual system. The supposition is that the various sensory manifestations of vision are expressions of the very same neural mechanism, simply acted upon in different ways. This approach is greatly in contrast to the older tendency to construct as many mechanisms as there are classifiable types of sensory end result.

B. STATEMENT OF PROBLEM

Intermittent stimulation will result in increased brightness within a certain range of circumstances without a corresponding increase in photic intensity. The increased brightness of intermittent illumination over continuous illumination of the same intensity is known as brightness enhancement or the Brücke-Bartley effect (1, 3, 7, 8). While the more conventional outlook might expect that whatever conditions produce brightness enhancement would improve visual acuity, there are more compelling reasons not to expect this simple relation. The conventional conclusion seems to rest either on a general confusion of intensity with brightness or arise from a general belief that whatever "improves" one sensory outcome should improve another.

Impingements play different roles when they are intermittent than when continuous. In the former, physiological processes leading to sensory end result must be repeatedly set up and terminated, whereas in the other, the end result once set up has only to be maintained. It is quite conceivable that whatever facilitates or improves the setting up of the end result in one case might not in the other. It is even quite possible to conceive of some conditions improving (raising or enhancing) brightness and at the same time impoverishing visual acuity. For example, in maintained brightness cross-sectional or lateral processes in the optic pathway are not involved in the full degree that they are in tests of visual acuity under intermittent stimulation. In maintained brightness, optic pathway activity is more or less uniformly distributed in time. Certain ranges of intermittent impingement tend to utilize the whole optic pathway for brief periods and little or none between photic inputs.

In the experiments to be reported, trains of photic pulses were employed. Both rate of intermittency and pulse-to-cycle fraction were manipulated. Individual elements in the trains varied from 5.2 to 187.5 msec. Some conditions of intermittency selected were such as to require resolution of target detail in fields in which brightness enhancement existed and others in which it did not exist. Specifically, acuity was studied under three conditions of continuous illumination of indefinite duration and eight conditions of intermittent illumination. In four cases, the photic pulse occupied one-fourth of the cycle (i.e., PCF .25) and was made intermittent at rates of 4, 10, 24, and 48 cps, and in the other four the PCF was .75. In both, the same frequencies were used.

C. METHOD

1. *Apparatus*

The investigation was conducted in a room 8 × 30 feet. At one end of the room *O* was seated at a chinrest fixed to a table supporting an upright black screen 30 × 40 inches. He viewed monocularly (right eye) through a peephole in the screen. At the other end of the room a lamphouse was suspended from an overhead track calibrated so that the target could be placed at six-inch intervals within a range of 19.5 to 26.0 feet from *O*. The side of the lamphouse facing *O* was made of opal glass and the transilluminated targets were placed against this surface.

The targets themselves consisted of a series of Snellen type E's centered on translucent sheets 5.25 inches square. Photometry of the materials showed the E's to be .25 as luminous as their background.

The targets could be placed in four positions so that the E's appeared in their normal position or *right,* i.e., with the open side of the E to the right; *up, left* or *down*. During experimentation three levels of intensity were used for non-intermittent conditions. The highest level was 173.55 c/ft^2. The other two intensities were three- and four-fold reductions of this value:

57.85 or 43.38 units. The intermittent conditions used one of the two higher intensity levels depending upon the PCF.

Four rates of intermittency, namely 4, 10, 24, and 48 cps were employed in combination with PCF's of .25 and .75. Thus there were eight conditions of intermittency. For all intermittency conditions involving PCF .25, intensity was set at the highest value, namely 173.55 c/ft^2. The flux per cycle was therefore .25 × 173.55 or 43.38 units. For all PCF .75 observations, intensity was set at 57.85 c/ft^2 and the flux per cycle was .75 × 57.85 or 43.38 units. The flux per cycle was therefore the same for all intermittency conditions, and could be expected to produce the same Talbot level. And the total flux per cycle was the same for the intermittent targets as it was for the non-intermittent target illuminated at the lowest level of intensity ($1.0 \times 43.38 = 43.38$ units).

An episcotister provided intermittency. It was mounted just back of *O*'s screen in order that onset and termination of pulses be as abrupt as we could make them. The diameter of the episcotister disk was made large in order to provide open sectors with radial subtenses quite large in relation to the peephole. This also contributed to producing rather abrupt onsets and terminations to the target presentations.

The room was kept illuminated at classroom levels during experimentation. The target area was at right angles to the overhead lamps however and when unlighted except for room illumination measured about 0.3 c/ft^2. The left eye was occluded with an eye patch to preclude its illumination by the room source.

2. *Observers*

Several groups of students served as *O*s. One group was comprised of eight students from an advanced undergraduate laboratory class. All had some experience observing intermittent targets, were familiar with the procedures of measurement and its general purpose, but were ignorant of the specific assumptions being tested. They produced the curves labeled A in Figure 1. Two other groups were comprised of totally naive *O*s. There were five different individuals in each group. These samples produced curves B and C of Figure 1. The senior authors also served as *O*s and their data are plotted separately.

3. *Procedure*

O was seated and told that he would see a target through the peephole; that this target was an "E" like the one on the wall chart nearby. He was told the "E" would appear in random order at each of four positions which he could call "up," "down," "left," and "normal" or "right." He was instructed to respond as quickly as possible and to make a response each time even if the best he could do was to guess.

Threshold was found for the three conditions of non-intermittency and afterwards for the eight conditions of intermittency. The intermittent conditions were ordered at random. Various modifications of the method of limits were used throughout.

4. *Visual Acuity Criterion*

For samples A, B and the senior authors, ascending and descending values representing four out of five correct judgments were used. For sample C, a criterion of six correct of eight judgments was used for both ascending and descending directions. The values plotted in Figure 1 are the arithmetic means.

The limits for all samples except C were obtained by only varying distance

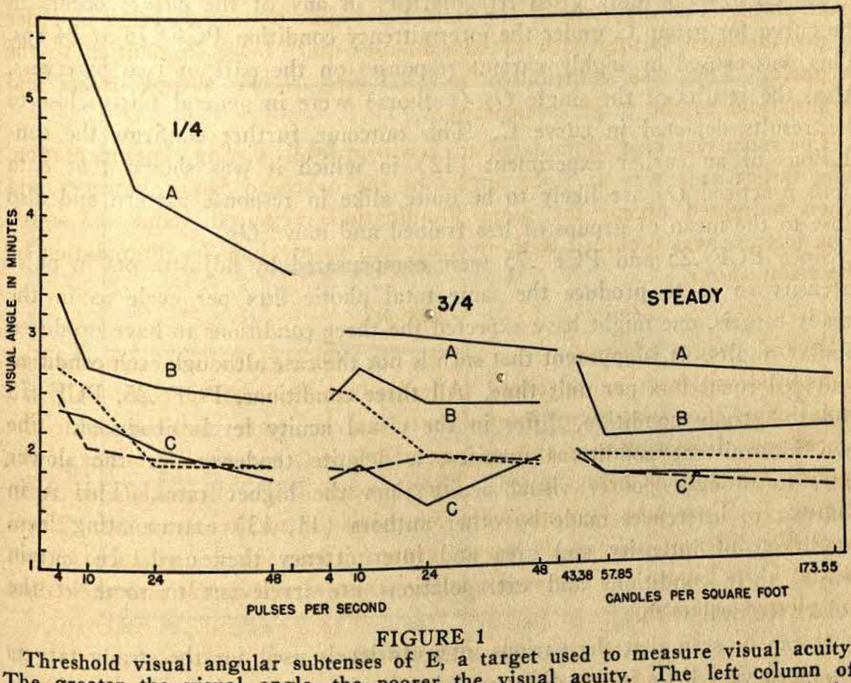


FIGURE 1

Threshold visual angular subtenses of E, a target used to measure visual acuity. The greater the visual angle, the poorer the visual acuity. The left column of curves are for intermittent target, PCF .25; the middle column, PCF .75; and the right column, steady illumination. Curves A, means for group of slightly trained observers; B and C, means for naive groups under two different procedures. Dashed curves, Nelson; dotted, Bartley.

around the range established by *O*'s first trial. This enabled all values to be obtained within a relatively short span of time, usually about one-half hour actual observation time, and limited the effects of practice. For sample C however, the whole range of target subtenses available was used in the presentation series until the criterion of 75 per cent correct judgments was reached. This was very time-consuming in that 4-10 hours was required to collect the experiment data for a single *O*. Hence C represents data collected over a number of days. The essential features of results obtained by this means were: (*a*) generally lowered thresholds; (*b*) greater variability between *O*s, perhaps because of actual day to day variability in acuity; but (*c*) a mean curve having the over-all characteristics of the other two groups.

D. Results

In general, the curves show an increasing visual acuity as the rate of intermittency rose from four per second to 48 per second, which is somewhat below CFF. The only gross irregularities in any of the curves occurs in the curve for group C under the intermittency condition PCF .75 at 24 cps. This was caused by highly variant responses on the part of two observers. Also, the results of the single *O*s (authors) were in general fairly close to the results depicted in curve C. This outcome further confirms the conclusions of an earlier experiment (12) in which it was shown that data from practiced *O*s are likely to be quite alike in response pattern and also close to the mean of groups of less trained and naïve *O*s.

Since PCF .25 and PCF .75 were compensated by adjustments in pulse intensity so as to produce the same total photic flux per cycle as in the steady targets, one might have expected the three conditions to have produced similar results. It is apparent that such is not the case although each condition provided equal flux per unit time. All three conditions, PCF .25, PCF .75 and the steady condition, differ in the visual acuity levels produced. The two intermittent conditions manifest a definite tendency for the slower rates to produce poorer visual acuity than the higher rates. This is in contrast to inferences made by other authors (11, 13) extrapolating from conditions of intensity and area and intermittency they used. In certain senses, their conditions and extrapolations are irrelevant to some of the points we deal with.

It can be seen that the various intensity levels used for the steady targets produced very little if any differences in visual acuity at the two upper levels while the lowest level is somewhat poorer. Since it was the two upper levels of intensity that were used to compensate for PCF in the intermittent target conditions, one can say that the factor of intensity as such probably didn't produce much if any of the difference found between the two intermittency conditions shown in Figure 1. In fact quite the opposite is true for lower rates of intermittency since the *higher intensity* (PCF .25) produced the *poorest acuity*. In order to understand this paradoxical outcome it is necessary to consult brightness curves related to these conditions. Such curves will allow one to state a relation between brightness produced by "enhancement" conditions and brightness produced by intensity conditions *and* visual acuity under the two conditions.

Brightness readings were obtained under the conditions used to obtain visual acuity readings. Figure 2 indicates the results. The left curve shows the relation of brightness to pulse frequency for a PCF of .25 for experienced

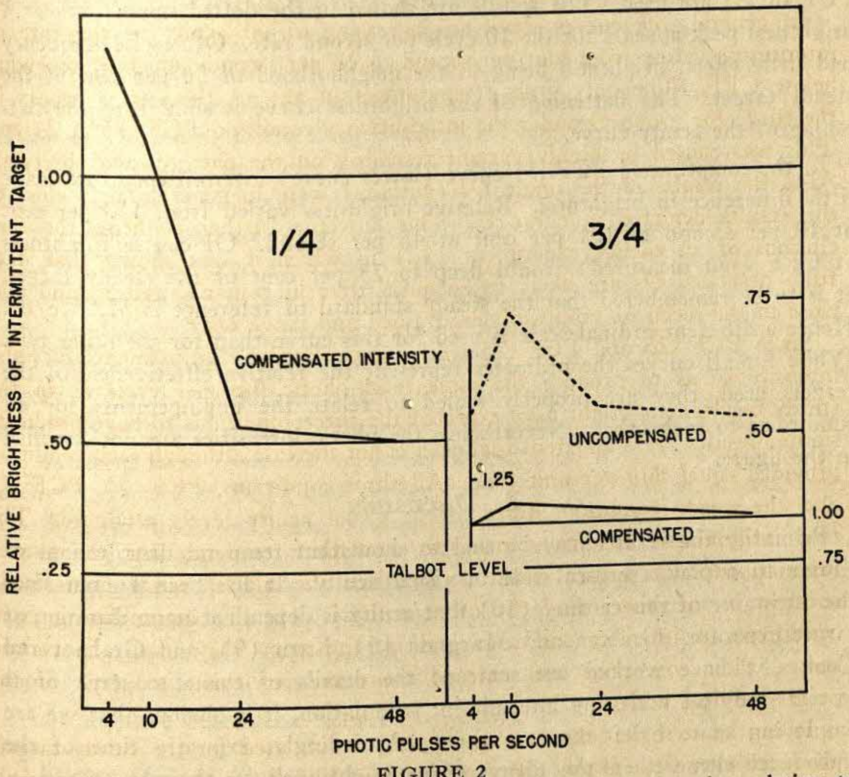


FIGURE 2

Curves for brightness of intermittent target at various frequencies relative to brightness of steady target. Left curves, PCF .25; right curves, PCF .75. Compensation means that PCF is multiplied by intensity to equal same total flux per cycle as in steady target.

observer N. It will be remembered that when visual acuity was measured under these conditions, PCF was compensated by raising intensity. This was done here, also. Hence an intensity of 173.55 c/ft^2 was used, and consequently the ordinate in Figure 2 uses this value as a reference. It will be noted that as pulse frequency is reduced, brightness increases quite markedly. Values for all frequencies are above Talbot level, and the values for both four and 10 per second are within the enhancement range. That is, *the two conditions producing the poorest acuity of all were those in which the brightness value of the intermittent was greater than that of steady illumination of the same intensity.*

The right curves have to do with a PCF of .75. In the case where intensity was not compensated, the same intensity and ordinal scale as applied to the PCF of .25 are used. The results are shown in the dotted upper curve. A brightness peak appears for the 10 cycle per second rate. Otherwise frequency had little effect, brightness being in the neighborhood of 50 per cent of the steady target. The flattening of the brightness curve is similar to the flattening of the acuity curve.

In the compensated PCF .75 curve (lower curve) intermittency rate made little difference in brightness. Relative brightness varied from 110 per cent at 10 per second to 101 per cent at 48 per second. Of course brightness at CFF (not measured) would drop to 75 per cent of the steady target. It is to be remembered that the steady standard of reference is 57.85 c/ft^2. Hence a different ordinal scale is used for this curve than for the other two. While for all curves the ordinates represent the relative effectiveness of the targets used, they are properly scaled to relate the impingements for all conditions to each other. Nevertheless the photic intensities are not specified in the figure.

E. Discussion

Primarily the results may be said to show that temporal distributions of illumination play a critical role in visual acuity. It has been known since the beginning of this century (10) that acuity is dependent upon duration of target exposure. Graham and Margaria (6), Karn (9), and Graham and Cook (5) have worked out some of the details of this discovery. Since in our study we are using intermittent stimulation, it is obvious that we are employing quite brief exposure times. Accordingly, exposure time of the individual elements in the photic trains might well be thought to play a decisive role in our visual acuity findings.

There are two reasons for not thinking this is the case. First, there is reason to think that the intensities used in the present experiment prohibit dependency of visual acuity upon single pulse duration (5). Second, and more decisive, is the fact that for PCF .25 durations of single pulses were at maximum where acuity was a minimum. If the duration of single pulses were the basis of the phenomenon then visual acuity should have had the opposite relation to rate.

From this one must infer that the results are due to intermittency per se, i.e., to the net effect of impingements which through their interruptions fail to best maintain conditions necessary for sensory resolution.[2]

The information and conclusions summarized in the alternation of response theory describe the nature of activity occurring in the optic pathway (1, 2, 3). This is a neurophysiological theory. The nervous events dealt with are those taking place within the optic pathway as a whole.

This pathway from eye to brain may be considered to consist in a fixed number of parallel channels. In the optic nerve segment of this pathway, the principle is most clearly recognizable since each nerve fibre is one of these channels. According to the theory, stimuli activate all, many, or few of these channels depending upon the strength of impingement, and certain temporal factors such as impingement duration, repetition, etc. Brief maximal inputs to the pathway activate all or nearly all channels simultaneously. Channels activated together tend to recover together, although there may be a slight variation in recovery times. When stimulation is made intermittent at a rate matching the mean recovery rate of the channels, the maximal number of channels stay in phase with one another. The alternation of response theory states that brightness varies in accord with the number of channels simultaneously activated. Brightness enhancement (the greater effectiveness of intermittent as compared to steady illumination of the same intensity) occurs because low PCFs and slower rates are those conditions producing maximal synchrony and hence maximal bursts of discharge.

Other conditions such as PCF .75, stimulations at fusion, and in particular, steady illumination distribute discharge. After a brief period of steady illumination there will be as many channels going into action as going into recovery phase in any given instant. This amounts to saying that uniform, prolonged impingements will elicit uniform over-all response in the pathway.

The negative correlation between brightness and acuity under conditions producing brightness enhancement lead one to surmise that, so far as longitudinal nervous events are concerned, visual acuity is best when discharge in the optic pathway is uniform and worst when discharge is characterized by maximal variation in activity.

---

[2] Independently, C. M. Bourassa has obtained results essentially similar to ours. Bourassa measured the necessary separation between two reflecting surfaces. Hence our data are not accounted for by something unique about our targets. It will be remembered that our target E was partially transilluminated. Bourassa's study will be published soon.

One may thus conclude that lateral nervous factors are not the only limiting features for visual acuity. This is not to deny the role of these factors in ordinary visual acuity situations but only to point out that visual acuity under intermittent conditions is quite largely determined by the temporal distribution of nervous events in the optic pathway.

F. SUMMARY

As brightness induced by intermittent stimulation increases, visual acuity worsens. As brightness induced by intensity increases, visual acuity improves. Thus brightness induced by these two means does not have summative but counter effects. This generalization refers only to ranges of intensity, etc. producing brightness enhancement and the intermediate range (1).

The results suggest the manner of operation of the optic pathway in determining acuity levels under intermittent stimulation. More specifically the results suggest that the "Alternation of response theory" has predictive implications for visual acuity.

REFERENCES

1. BARTLEY, S. H. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Opthalmol.*, Pt. II, 1958, **60**, 775-791.
2. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. *In* Jung, R., & Kornhauser, H. (*Eds.*), *The Visual System: Neurophysiology and Psychophysics*. Heidelberg, Germany: Springer-Verlag, 1961.
3. BARTLEY, S. H., & NELSON, T. M. Some relations between sensory end results and neural activity in the optic pathway. *J. of Psychol.*, 1962 (in press).
4. GERATHWOHL, S. J., & TAYLOR, W. F. The effect of intermittent light on readability of printed matter under conditions of decreasing contrast. *J. Exp. Psychol.*, 1953, **46**, 278-282.
5. GRAHAM, C. H., & COOK, C. Visual acuity as a function of intensity and exposure time. *Am. J. Psychol.*, 1937, **49**, 654-661.
6. GRAHAM, C. H., & MARGARIA, R. Area and the intensity time relation in the peripheral retina. *Am. J. Physiol.*, 1935, **113**, 299-305.
7. GRÜSSER, O.-J., & CREUTZFELDT, O. Eine neurophysiologische Grundlage des Brücke-Bartley-Effektes: Maxima der Impulsfrequenz retinaler und corticaler Neurone bei Flimmerlicht mittlerer Frequenzen. *Pflüg Arch.*, 1957, **263**, 668-681.
8. HALSTEAD, W. C. A note on the Bartley Effect in the estimation of equivalent brightness. *J. Exp. Psychol.*, 1941, **28**, 524-528.
9. KARN, H. W. Area and intensity-time relation in the fovea. *J. Gen. Psychol.*, 1936, **14**, 360-369.
10. MCDOUGALL, W. The variation of the intensity of visual sensation with duration of the stimulus. *Brit. J. Psychol.*, 1904, **1**, 151-189.

11. Nachmais, J. Brightness and visual acuity with intermittent illumination. *J. Opt. Soc. Amer.*, 1958, **48**, 726-730.
12. Nelson, T. M., Bartley, S. H., & DeHardt, D. C. A comparison of variability of three sorts of observers in a sensory experiment. *J. of Psychol.*, 1960, **49**, 3-11.
13. Senders, V. L. Visual resolution with periodically interrupted light. *J. Exp. Psychol.*, 1949, **39**, 453-465.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 165-168.

TASK DIFFICULTY, ANXIETY LEVEL AND ABILITY LEVEL AS FACTORS AFFECTING PERFORMANCE IN A VERBAL LEARNING SITUATION*[1]

Tufts University

BERNARD W. HARLESTON

A. PURPOSE

In an earlier experiment (Harleston and Cunningham, 1961) an attempt was made to investigate the effects of manifest anxiety on verbal learning as a function of the difficulty level of the learning task. Independent groups of low anxious (LA) and high anxious (HA) females and males learned one of four lists of 12 paired-associate nonsense syllables. The lists differed in terms of the relationship of the association value of the stimulus items to the response items. It was hypothesized that the difference in performance between LA and HA *S*s would increase as the difficulty level increased. No statistically significant main effects due to anxiety level were found. However, it was observed that the list producing the greatest difference between LA and HA *S*s differed for the females and the males. For the former, the low-low list, consisting of stimuli and responses of low association value, produced the greatest difference. In the case of the males, the low-high list produced the greatest differences. Since the females learned the practice list of meaningful word pairs significantly faster than the males, the *ad hoc* hypothesis was formulated that the operation of anxiety effects depends upon both task difficulty level and ability level.

The present report represents an attempt to test this hypothesis. In this approach, the original experiment was replicated so that with the combined data of the two experiments, three ability level groups, as measured by performance on the practice test, would be available for each list at each anxiety level. The experimental design involved is a 2×2×2×3 four-way analysis involving, respectively, stimulus association value, response association value, anxiety level, and ability level.

---

* Received in the Editorial Office on October 14, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was supported in part by a grant (M-3597(A)) from the National Institute of Mental Health.

B. METHOD

1. *Subjects*

Ninety-six females and 96 males served as *S*s. Half of each sex had scores of eight or less (LA) on the 50-item form of the Taylor Scale of Manifest Anxiety; the remaining *S*s had anxiety scale scores of 23 or above (HA).

2. *Procedure*

The learning materials and procedures have been described in detail elsewhere (Harleston and Cunningham, 1961). In brief, *S* was given eight trials on a practice list of 12 pairs of meaningful, five-letter words. This was followed by the presentation of the appropriate experimental list. Three different orders of the practice list and four different orders of each experimental list were used to minimize serial effects.

C. RESULTS

The mean number of correct responses made by the various treatment groups is presented in Table 1. These data were evaluated for statistical significance by a four-way analysis of variance. Significant main effects were obtained for stimulus association value ($F = 12.40$; $df = 1/168$; $P = <.001$), response association value ($F = 38.44$; $df = 1/168$; $P = <.001$), and ability level ($F = 32.20$; $df = 2/168$; $P = <.001$). Significantly more correct responses were given when the stimulus members were of high association value than when they were of low association value. Similarly, more correct responses were made when the response items were of high association value than when they were of low association value. With respect to ability level, the fast learners produced significantly more correct responses than the moderate learners ($P = <.001$), and the moderate learners produced significantly more correct responses than the slow learners ($P =$

TABLE 1
MEAN NUMBER OF CORRECT RESPONSES ON THE VARIOUS LISTS FOR HIGH AND LOW ANXIOUS *S*S OF EACH ABILITY LEVEL

	Slow learners		Moderate learners		Fast learners	
List	LA	HA	LA	HA	LA	HA
H-H	159.50	156.63	201.38	190.13	228.75	213.13
L-H	138.25	90.38	161.13	136.13	225.88	191.50
H-L	90.38	79.00	129.88	150.63	183.25	182.13
L-L	87.88	85.00	103.75	119.88	157.38	131.13

<.001). Neither anxiety level nor any of the interactions produced statistically significant effects.

Because the ability factor was demonstrated to be highly significant, its effect on learning was explored further. Performance on each list was evaluated by an anxiety by ability by trial-blocks (six blocks of five trials) repeated measures design. The *F* ratios and associated *P* values of the various sources of variance for each list are presented in Table 2. Significant effects due to ability level, trial-blocks and the interaction of ability level and trial-blocks were found for each list. This latter effect means that on each list the rates of learning varied directly as a function of ability level. Additional significant effects for particular lists, and their interpretation, are as follows: On the high-high list, the significant trial-blocks by ability level by anxiety interaction reflects the finding that the slow learning LA *S*s had a faster learning rate than the slow learning HA *S*s for the first 15 trials, and a slower rate for the last 15 trials. On the low-high list, the significant anxiety effect results from the superior performance of the LA *S*s over all trials. The significant trial-blocks by anxiety interaction on this same list reflects the faster rate of learning manifested by the LA *S*s.

TABLE 2

SIGNIFICANT *F* RATIOS AND *P* VALUES FOR THE COMPONENTS OF ABILITY LEVEL BY ANXIETY LEVELS BY TRIAL BLOCKS ANALYSES OF VARIANCE OF EACH LIST

Source of variance	Lists							
	H-H		L-H		H-L		L-L	
	F	*P*	*F*	*P*	*F*	*P*	*F*	*P*
Ability (Ab)	3.37	.05	15.27	.001	16.09	.001	5.10	.01
Anxiety (A)	—	—	6.41	.03	—	—	—	—
Ab × A	—	—	—	—	—	—	—	—
Trial-blocks (TB)	276.14	.0001	295.96	.0001	249.23	.0001	205.96	.0001
TB × Ab	2.45	.01	4.44	.001	4.83	.001	3.12	.001
TB × A	—	—	5.08	.001	—	—	—	—
TB × Ab × A	2.03	.03	—	—	—	—	—	—

D. DISCUSSION

The hypothesis that the effects of anxiety level on performance are a function of task difficulty level and ability level did not receive support. While both difficulty level and ability level emerged as highly significant, performance was generally not sensitive to effects due to anxiety level. On only one list were differences due to anxiety level reliably obtained. And in this instance anxiety level did not interact with ability level. While the significant trial blocks by ability level by anxiety level interaction on the high-high list does involve the low ability groups and the reversals in learning rates as a function of anxiety level, this effect seems too limited to be interpreted as direct support. Thus there is little basis for the view that the effects of anxiety level depend upon the ability level of the *S*s.

The highly sensitive role ability level played in this experiment assumes added significance in view of increasing emphasis on individual differences and learning. Recently Noble (1961) has argued that considerably more attention should be given to the development of measures of individual differences and to the identification of major experimental variables relevant to individual differences. The data of this experiment demonstrate that the operational procedure of defining ability in terms of a small sample of performance with meaningful material is an inexpensive and effective approach to the measurement of individual differences.

E. SUMMARY

The interactive effects of Taylor scale anxiety level, ability level and difficulty level were investigated with a paired-associates learning task. While ability level and task difficulty level emerged as statistically significant effects, anxiety level failed to appear as consistently significant. Further, the hypothesis that the effects of anxiety level on performance are a function of task difficulty level and ability level did not receive support. The implications of the sensitive role of ability level were discussed.

REFERENCES

1. HARLESTON, B. W., & CUNNINGHAM, S. M. Task difficulty and anxiety level as factors affecting performance in a verbal learning situation. *J. of Psychol.*, 1961, **52**, 77-86.

2. NOBLE, C. E. Verbal learning and individual differences. *In* Cofer, C. (*Ed.*), *Verbal Learning and Verbal Behavior*. New York: McGraw-Hill, 1961.

Department of Psychology
Tufts University
Medford 55, Massachusetts

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 169-173.

A SURVEY OF SOCIAL PSYCHOLOGY COURSES IN TEACHER TRAINING INSTITUTIONS*

Newark State College

ROBERT HOWARD ROTH

A. PROBLEM

This study is one of a series of considerations of the teaching of social psychology courses in American institutions of higher learning.

Anyone who works in teacher training institutions is probably aware that results of social psychological theory and research are quoted and presented rather often within the various professional courses of the pre-service curricula. Specific courses in the discipline generating this work, however, have not yet been studied to determine the rationale behind such courses, the *how*, the *why* and the *what* of them. The purpose of the present study is to begin this task of finding out what is currently being done in such courses together with some related information from and about those who teach the courses.

Teacher training has already found social psychology to be a most important discipline. It is possible that a good social psychology course, integrated and consonant with the purposes of pre-service curricula would be a valuable and important part of such programs. Before a prescription, however, an adequate diagnosis is in order. It is hoped that this study will be a contribution in that direction.

B. METHOD

A sample of 120 teacher training institutions was drawn from the membership list of 534 colleges of The American Association of Colleges for Teacher Education.

A questionnaire and a covering letter, developed originally for use in a prior study of social psychology courses in liberal arts colleges,[1] was mailed to members of the sample. The questionnaire consisted of 23 items dealing with information in four main areas of interest: (*a*) the administrative details of the course; (*b*) the subject matter content of the course; (*c*) teaching methods

---

* Received in the Editorial Office on September 24, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Robert H. Roth, "A Survey of the Undergraduate Course in Social Psychology," *The Journal of Social Psychology* (in press).

and related course procedures; and (*d*) the course instructors' background and opinions on various relevant topics dealing with social psychology.

C. Results and Discussion

1. *Administrative Data*

Returns were received from 70 institutions (58 per cent of the sample). Of these returns, 36 institutions (30 per cent of the sample) reported that they offered no social psychology course. Returns suitable for tabulation were received from 34 institutions (28 per cent of the sample).

Most courses were reported as running for 15 weeks, with an average of three meetings per week, each class meeting lasting from 50 to 60 minutes. Average number of students per class was 31.

Prerequisites for admission to the course were required by all except four of the colleges reporting. General Psychology was most generally required (22 institutions) with General Sociology the sole requirement of two colleges.

2. *Content of Course*

Instructors indicated mainly three different definitions of social psychology. *The Study of the Individual as Influenced by Social Factors* was the working definition used by 50 per cent of the instructors. *The Study of the Relation of Individuals and Groups* was reported by 24 per cent of the instructors, and *The Study of Social Interaction* was favored by 21 per cent of the instructors. As reported in a previous study[2] there proved to be a difference in the way the single most often indicated definition, *The Study of the Individual as Influenced by Social Factors*, was interpreted. Instructors who favored a sociological emphasis stressed the *Social Factors* part of the definition as the main point of focus in understanding human behavior. Those who favored a psychological focus stressed the *Individual* as the main point of focus for the unravelling of the roots of behavior.

Textbooks regularly assigned by classroom instructors were ranked as follows: 1. Sargent and Williamson, 2. SPSSI *Readings in Social Psychology*, 3.5. Britt, 3.5. Lindesmith and Strauss, 6. R. Dewey and Humber, 6. Klineberg, 6. Newcomb. Additional required readings were widely varied, with the most often mentioned books by Riesman, C. W. Mills, Festinger, M. Mead, and in the *Handbook of Social Psychology*.

Rank order of topics covered in the course was: 2. Attitudes, 2. Leadership, 2. Prejudice and ethnic relation, 5. Communication, 5. Propaganda, 5. Socialization and personality development, 7. Group development and behavior, 8.5. Social Perception, 8.5 Public opinion, 10. Collective behavior, 11. Basic motives, drives or instincts. Topics ranked for amount of emphasis given in class were: 1. Socialization and personality development, 2.5. Attitudes, 2.5. Culture and personality, 4. Communication, 5. Prejudice and ethnic relations, 6. Group development and behavior, 7. Propaganda, 8. Social perception.

---

[2] Op. cit.

Theoretical positions used as a basis of class instruction were ranked: 1. Role Theory, 2. Field Theory, 3. Stimulus-response contiguity and reinforcement Theory, 4. Psychoanalytic Theory, 5. Cognitive Theory. The theoretical positions used most often were ranked: 1. Role Theory, 2.5. Field Theory, 2.5. Stimulus-response contiguity and reinforcement Theory, 4. Psychoanalytic Theory, 5. Cognitive Theory.

Various methods of organizing the course were reported. Rank order of these methods was: 1. Instructor follows the organization of the textbook he has chosen, 2. Course begins with focus on psychological variables, moves to focus on group variables, and concludes with focus on macro-social variables, 3. Focus is on several basic topics which are studied intensively, 4. Focus is entirely on socialization processes at the group level.

3. *Teaching Methods*

Rank order of those teaching methods frequently used was: Lectures, 2. Class discussions, 3. Tests, 4. Student oral reports. Rare use or no use at all was most often reported for: 1. Recordings played in class, 2.5 Field work, 2.5. Laboratory work, 4.5. Films shown in class, 4.5. Role-playing, 6. Small group work (21 of the 34 institutions report rare use or no use at all of this teaching method!), 7. "Buzz" sessions.

Research techniques practiced in class were ranked: 1. Observation of social behavior, 2.5. Experiments, 2.5. Sociometric measurement, 4. Attitude measurement, 5. Interviewing, 6. Statistics.

Kinds of small group work most often used, when used at all, were: 1. Discussion groups, 2. Project teams or groups, 3. Role-playing groups, 4. Group dynamics training groups.

One third of the instructors reported that students were given little or no responsibility in either choosing topics for the course or in leading class discussions. One half or more of the instructors reported that students were given little or no responsibility for determining assignments, preparing class examinations, or determining the grades.

Instructors reported that they determined final grades by placing the following average percentages of weight on various aspects of the course: test results 60 per cent, student participation in discussion 15 per cent, written reports 15 per cent, oral reports 5 per cent, small group work 5 per cent.

The impression that is reinforced by a study of the data indicates that teachers of social psychology courses for future teachers follow, on the whole, traditional procedures in their teaching methods. The very material that they teach, however, is laden with hints of new directions in man's understanding of his own learnings within the human groups of which he is a part. Teachers of social psychology do not seem to be transforming their own teaching practices as a result of the insights which theories and research in social psychology have achieved. Perhaps it is not irresponsible to hope for the day when the very subject which has contributed so greatly to more mature explanations of man's behavior patterns will be taught in a way that puts into practice some of its own discoveries!

4. *Instructors' Background and Opinions*

In addition to social psychology courses, rank order of courses taught most often by instructors was: 1. General Psychology, 2. Personality, 3. Abnormal Psychology, 5.5. Child Psychology, 5.5. Adolescent Psychology, 5.5. General Sociology, 5.5. Mental Health.

Instructors had taught social psychology courses for an average of eight years apiece. There were 27 Ph.D.s and four Ed.D.s among this group of 34 instructors. The approximate distribution of percentage of professional time spent on various activities was: Teaching 70 per cent, Research 20 per cent, Administrative duties 2 per cent, Lecturing (outside classes) 2 per cent, Consultation 2 per cent, Therapy 2 per cent.

Major subjects in instructors' undergraduate programs were ranked: 1. Psychology, 2. English, 3. Biological Sciences, 4. Sociology. Major subjects in Master's degree programs were: 1. Psychology, 2. Education, 3. Sociology, 4. Social Psychology. Major subjects in doctoral programs were: 1. Psychology, 2. Sociology, 3. Education, 4.5 Philosophy, 4.5 Social Psychology.

Instructors indicated that within the organization of instructional areas social psychology belongs most appropriately in the Psychology department (N = 19), or as a joint offering of both the Psychology and Sociology departments (N = 11), or in the Sociology department (N = 3).

The rank order for the names of those whose work had influenced the instructors' own thinking about social psychology was: 1. G. Allport, 2. M. Mead, 3. G. Murphy, 4.5. Lewin, 4.5. T. Newcomb, 6. M. Sherif, 7. O. Klineberg, 8. S. Freud, 9. S. Asch, 10. K. Horney, 11. J. Moreno, 13.5. J.

Dollard, 13.5. E. Fromm, 13.5. G. H. Mead, 13.5. H. S. Sullivan. The rank order for the names of those whose work had the *greatest* influence on the instructors' thinking about social psychology was: 1. G. Allport, 2. K. Lewin, 3. M. Mead, 4. S. Freud, 5. T. Newcomb, 6.5. O. Klineberg, 6.5. M. Sherif, 8.5. S. Asch, 8.5. G. H. Mead, 10. K. Horney, 11. E. Fromm.

The crucial areas of future exploration for social psychology were ranked: 1. Group dynamics, 2. Social perception, 3. Prejudice, 4. General theory construction, 5. International relations. The crucial areas of application for the current results of social psychology's investigations were ranked by instructors as: 1. Prejudice and intergroup relations, 2. Increasing the effectiveness of groups, 3. Problems of international relations, 4. Exposure of propaganda techniques, 5. Prevention of juvenile delinquency and crime.

Perhaps it is significant that the instructors did not specifically mention educational institutions as sectors of the modern world where the results of social psychology should be applied, though it might be said that this is at least implied in some of the highest ranking responses. Yet, might it not be that this relates to the rather negligible use of social psychological insights in the instructors' own teaching practices, in so far as these can be inferred from the questionnaire responses?

It would seem that there exist areas where considerable improvement in the presentation of social psychology to future teachers is possible and desirable. If we wish our students to become better teachers by using the results of social psychology, perhaps we can best teach them the desirability and efficacy of this application of behavioral science theory and research by practicing it ourselves.

D. SUMMARY

Results of a survey of the teaching of social psychology courses in teacher training institutions are presented and discussed. Inferences are drawn from the data regarding the possible lack of sufficient practice of what is presented in the subject matter of the course. The importance of practicing what one professes is nowhere more crucial than in the educating of future teachers. Social psychology, with so much to say about the genesis of behavior, ought to be taught in such a manner that its content is illuminated by the classroom environment within which this material is presented, discussed and, hopefully, learned.

Newark State College
Morris Avenue
Union, New Jersey

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 175-182.

A LOGICAL METHOD FOR MAKING A CLASSIFICATION OF EMOTIONS, USING WILHELM WUNDT'S THEORY OF EMOTION FORMATION*

Department of Philosophy, Rutgers University College

WILLIAM L. ROSENSOHN

A. WUNDT'S THEORY

The German philosopher-psychologist Wilhelm Wundt held the following theory about the formation of emotions. The basic elements of consciousness, he believed, are simple sensations and feelings. Out of these arise the psychic compounds, to wit, ideas or perceptions, composite feelings, emotions and volitions.

Using experimental techniques, Wundt succeeded in isolating three pairs of simple feelings which he regarded as ultimate. They were the feelings of: (*a*) pleasure and displeasure; (*b*) strain and relaxation; and (*c*) excitement and quiescence (or depression). The metronome experiment Wundt describes in his *Introduction to Psychology* (1) illustrates his discovery.

As the following passage from the same work (1) makes clear, the emotions themselves are compounds of these simple feelings:

> An important characteristic of feelings consists lastly in the fact that they combine themselves into an affective process, which as a rule is joined to an ideational process. A temporal process of this kind with an affective and ideational content that changes but nevertheless is joined together, we call an emotion, or with less intensity and a more lasting nature of the feelings, a disposition.

Volitional processes are marked out from the emotions by having: (*a*) ideas with more or less strong feeling-tone (the motives of volition); and (*b*) feelings of activity which accompany the end-stage. The distinction made here seems significant since it focusses attention upon the *relatively weak ideational content* of ordinary emotions, as conceived by Wundt.

B. THE PROBLEM OF CLASSIFICATION

Given this analysis, Wundt tried to classify the emotions but without success. In his volume, *The Outlines of Psychology* (2), he discusses and

---

* Received in the Editorial Office on September 25, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

rejects as unsatisfactory three separate classificatory schemes which may be briefly summarized.

1. The first scheme groups the emotions according to quality of feeling. Three main classes are distinguished: One, the subjectively distinguished emotions, whose chief primary forms are joy and sorrow with subforms, sadness, care, grief, fright and so on; two, the objectively distinguished emotions, whose primary forms are delight and displeasure with subforms, annoyance, resentment, anger, rage, *etc.*; and three, those referring to some future event, the primary forms of which are hope and fear with modifications, worry, anxiety and so on.

2. The second mode of classification divides the emotions according to intensity of feeling. Two classes of emotions are formed: strong and weak.

3. The third classifies as to mode of occurrence, giving the psychologist three classes. One includes the sudden, irruptive emotions. Examples are surprise, astonishment, fright, disappointment, rage. Another comprises the gradually arising. Examples are anxiety, doubt, care, mournfulness, expectation; in many cases, joy, anger, worry. The last includes intermittent, fluctuating, rising and falling emotions. Joy, anger, mournfulness come in waves.

In the opinion of Wundt classification 1 neglects the quantitative side of feelings while neither 2 nor 3 provides a fixed criterion for the psychological classification of the emotions in actual cases. Accordingly, at the conclusion of the discussion referred to above, Wundt proposes the following rule for successful classification:

> A classification to any degree exhaustive must subdivide such varying emotions as joy, anger, fear and anxiety into their subforms, according to their mode of occurrence, according to the intensity of component feelings, and finally according to their physical concomitants.

It is the contention of the writer of this paper that Wundt's rule ought to be amended. I specifically contend that a successful classification of emotions according to Wundt's theory of emotion-formation can be made provided that two fundamental classification decisions are executed. One decision, however, would result in eliminating Wundt's difficult first requirement, classification according to mode of occurrence. As used by Wundt, mode or form of occurrence refers to the way an emotion develops, as suddenly, irruptively, gradually, intermittently or in rising and falling waves.

The decision is that the classification scheme should be simple and noncomplete. I think no question of logic enters here. The English alphabet is theoretically incomplete but still sufficient to meet the needs of those who speak English. That is to say, it is limited by convention: by means of it we can reproduce only a limited series of sounds, the sounds of the English language.

Therefore, it is proposed that we eliminate three classes of emotions and thus, in effect, put aside, for purposes of classification, considerations of mode of occurrence.

First, leave out the more or less fixed dispositions as greediness or cowardice. Second, the immoderate emotions as lust, depression or dread, which figure so prominently in our lives and literature. These emotions require a separate classification. Last, the mixed or *complex* emotions, which are by far the most common. These *defy* classification, and it is here, I believe, that a classifier runs into insurmountable difficulty.

An actual emotion is never simple but a complex cluster. Actual rage is never isolated but clustered about by feelings (emotions) as various as fright, joy, sorrow, worry, anxiety, surprise, lust, contempt, self-hatred, pity, derision, liking, loathing, pride, shame, cowardice; we may nearly empty the vocabulary of emotions in describing actual rage.

The *second* decision is that the classification scheme should be useful but non-instructive. To put it another way, emotions should be defined in emotionally neutral terms. No emotion should be regarded as a subform of another but rather as related within a class. Thus, under this ruling it would be illegitimate to *define* jealousy in terms of hate or pride in terms of love. Instead, it is proposed that we adopt Wundt's classification according to (*a*) certain physical concomitants and (*b*) intensity of component feelings, both of which are emotionally neutral. (Whether or not such a scheme, if successful, is useful goes beyond the scope of this paper.)

A *logical* classification now appears possible.

C. CLASSIFICATION

In Chapter 13, "Emotions," (2) Wundt makes the following points:

> 1. Generally only a single affective tendency can be called primary for a particular emotion.
> 1.1. These primary affective elements are either feelings of pleasure or displeasure or feelings of strain or relaxation.

We shall, accordingly, formulate four categories of primary emotions: the primarily pleasurable, the primarily displeasurable, the primarily strained and the primarily relaxed, which we will represent respectively by the letters P, D, S and R.

Wundt further declares:

> 2. There are affective elements belonging to other dimensions than the primary which enter in as secondary elements.
> 2.1 Primarily strained or relaxed emotions may easily be associated with pleasurable or displeasurable feelings.
> 2.2 Feelings of pleasure or displeasure are *always*, according to circumstances, either exciting or depressing.

Thus, Wundt's theory of emotion formation allows us to introduce, for each primary emotion (D, P, S or R), one or two variables by means of which the various subforms of D, P, S or R may be arranged or ordered. The variable we choose will be *excitement* which we shall translate into the dichotomy, *quiet*, *arousing*. The first part of the dichotomy may be represented by the letter q, the second by the letter a. To define a serial order or continuum of emotions we shall introduce *standards*.

It should be noted here that we choose the single variable *excitement* rather than Wundt's pair of variables, *excitement* and *depression*, in accordance with our earlier decision in favor of simplicity and non-completeness wherever possible. By so doing we have limited ourselves to a classification of ordinary emotions *only*. In dealing with immoderate emotions the *two* variables required by Wundt's theory must be used. A classification of this type, with a tentative interpretation, will be discussed briefly in the last section of this paper.

Finally, it is proposed that we classify any emotion according to how we decide the question, "How is the emotion in question generated?" Only two ways are suggested by the writer who here is responsible for adding a new distinction (provided by the old Greek insight concerning the affinity or unity of opposites) to the Wundtian analysis. Either an emotion is generated by *overcoming* its primary feeling or it is generated by the primary feeling's *emergence* into full consciousness. When a primary feeling forming an emotion is overcome or cancelled out, we shall represent that by placing a minus sign in front of the letter representing the emotion; where the feeling is emergent, the letter representing the motion stands alone.

Thus we get the schema, presented in Table 1.

If F in Table 1 stands for any primary feeling, then Fqa and —Fqa are also formulas and are to be interpreted as *continuums* of emotion.

D. Interpretation

The problem of interpretation boils down to the question, "What standards ought to be adopted?" The answer given here may not satisfy the ex-

TABLE 1

CLASSIFICATION OF ORDINARY EMOTIONS WITHOUT INTERPRETATION

	F emergent			F overcome		
F	q	⟶	a	q	⟶	a
D	Dq		Da	—Dq		—Da
P	Pq		Pa	—Pq		—Pa
S	Sq		Sa	—Sq		—Sa
R	Rq		Ra	—Rq		—Ra

perimental psychologist, but I can offer no better one: It is to take the ordinary words in our language, the words which we use every day in describing our emotions, and fit them into our schema. Thus we may define each of our theoretic formulae by assigning to each a word out of the ordinary language.

A possible interpretation is given in Table 2.

TABLE 2

CLASSIFICATION OF ORDINARY EMOTIONS WITH TENTATIVE INTERPRETATION

	F emergent			F overcome		
F	q	⟶	a	q	⟶	a
D	Dislike		Hate	Content		Joy
P	Liking		Love	Regret		Sorrow
S	Uneasiness		Anxiety	Confidence		Hope
R	Calm		Boredom	Surprise		Fear

The eight continuums we obtain—Dqa, –Dqa, Pqa, . . . , Rqa—may also be interpreted by means of the vocabulary in Table 2. Thus, –Pqa stands for the Regret-Sorow continuum, and so on.

Having defined our standards in every-day language terms, we are now in a position to classify the remaining ordinary emotions within the eight continuums furnished us by Wundt's basic analysis.

Unfortunately for the classifier, emotions, it seems probable, are interpretations of feelings, and reduction of an emotion to a fixed quantity within a continuum appears an impossible task. Here again, however, we must decide in favor of simplicity and somehow let words, at least those which are fairly definite, be our guide. The words will be part of our ordinary speech.

To take an example: how shall we classify the non-standard emotion, *anger?* We shall have to ask ourselves, What is it to be angry? It is to have a feeling, of course. Of what kind? Here clinical and experimental evidence must help us to make our decision. Of displeasure, let us say. But then how much? We cannot go by physical effects. The effects will vary from occasion to occasion. Sometimes circumstances will permit us to "let it out." Sometimes circumstances will have us keep it inside. Then how should we measure anger? It seems as if anger should be reducable to a fixed quantity, or else how can we really recognize it? It ought to be distinguishable from "a little more than anger" or "a whole lot more," and from "a little less than anger" and "a whole lot less."

In recognizing anger I think we do something like this. I think we recognize it by "seeing" or placing it between what is less and what is more. Now what is less than anger? The word "aggravation" helps us. Aggravation is a little less than anger. And what is more? Here we have the word "rage." Is rage merely "being really angry"? I don't think by "rage" we mean being really angry. Whenever we are angry, we are really angry. Otherwise, it is sham anger, and that is something else. I think that this is the difference between being angry and in a rage: When you are able to talk or think, you are still only angry, though you are really angry. When you are in a rage, your whole body reacts. You have stopped thinking or talking, have descended to making inarticulate sounds or you may even be stricken dumb with rage. (Effects allow us here to distinguish rage from anger. The effects, however, will not enable us to *recognize* rage. Grief may cause in us the same phenomenal effects as rage did.)

An any rate, we see that it is possible to fix anger as a *quantity* of emotion between rage and mere aggravation. Next we ask, What is more even than rage? To get away from the answer "really raging," we must ask, What follows rage as rage followed anger? Of course, there are two senses of "follow." One sense refers to a temporal sequence. The other sense refers to what transcends rage as an emotional quantity or value. In this latter sense, I say that sorrow follows rage. Finally, what is less even than aggravation? Why not regret? When we merely regret something, we can tell ourselves, "No use getting aggravated over that." But we do regret. So we see, with the aid of ordinary words, that simple anger is a kind of *mean* or proportional between regret and sorrow, where regret is quiet and sorrow exciting. Of course, we never feel just one emotion, as simple anger. Still, we *can see* that anger, though accompanied by other emotions, may be assigned to the regret-sorrow continuum, –Pqa, at a point between aggravation and rage.

According to this analysis, one should be able to assign every one of the ordinary emotions to a continuum, and to one only, and further be able, if there are words available, to fix an emotion between two others which are not standards but represent, as the standards do, more and less within the continuum.

I realize this is an ambitious program but I do not see why it cannot be done, given some kind of scheme based on Wundt's theory of emotion-formation. For example, using the method of classification and of interpretation already suggested, we are able to make the classification given in Table 3.

TABLE 3

CLASSIFICATION OF TEN ORDINARY (NON-STANDARD) EMOTIONS

Emotion	What is less than . . .	What is more than . . .	Classification
Curiosity	Being puzzled	Desire to know	Sqa "Uneasiness- — Anxiety"
Guilt	Desire to know	Remorse	
Shame	Remorse	Feeling worthless	
Feeling worthless	Shame	General anxiety	
Surprise	Expectation	Consternation.	—Rqa "Expectation- Fear"
Consternation	Surprise	Worry	
Worry	Consternation	Alarm	
Jealousy	Resentment	Hate	Dqa "Dislike-Hate"
Envy	Disappointment	Feeling of loss	—Pqa "Regret-Sorrow"
Anger	Aggravation	Rage	

Finally, even the *immoderate* (extreme) emotions, which go from extreme excitement to extreme depression, can be classified in terms of Wundt's theory of emotion formation, provided the variables, *excitement* and *depression,* originally suggested by Wundt (2), are used in place of *quiet* and *arousing.* Thus, using the letters e and d to represent respectively *exciting* and *depressing*, we get the results given in Table 4.

Classification is, of course, a way of looking, a way of grouping phenomena. As such it is subject to the classifier's own limitations and views and thus needs to be criticized as well as defended. Certainly one may ask why the writer supposes that guilt belongs to the uneasiness-anxiety continuum and not to expectation-fear, or why he assigns worry to the latter continuum and

TABLE 4

CLASSIFICATION OF IMMODERATE EMOTIONS WITH TENTATIVE INTERPRETATION

F	F emergent		F overcome	
	e	d	e	d
D	Extreme aversion	Nausea	Elation	Depression
P	Lust	Satiety	Grief	Melancholia
S	Dread	Horror	Euphoria	Apathy
R	Ennui	Utter weariness	Panic	Despair

not to the former. Further, how sure can he be that the emotion of anger ends in sorrow and not in hate? Moreover, are anger and envy, as he supposes, primarily displeasurable emotions or are they primarily strained? And, remembering Othello, ought we to think of jealousy in terms of hate rather than in terms of love? Questions such as these cannot be settled by logic or linguistic analysis. Only acute observation and/or experiment can corroborate our guesses.

REFERENCES

1. WUNDT, W. An Introduction to Psychology. London: George Allen & Co., Ltd., 1912. Pp. 43-63.
2. ———. Outlines of Psychology. 3rd, revised English edition. New York: G. E. Stechert & Co., 1907. Pp. 16-30, 31-40, 83-101, 176-181 and 188-203.

Department of Philosophy
Rutgers University College
New Brunswick, N.J.

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 183-198.

ROLE CONFLICT OF PUBLIC SCHOOL TEACHERS*[1]

Oklahoma State University and University of Missouri

J. PASCHAL TWYMAN AND BRUCE J. BIDDLE

A. INTRODUCTION

The role of the public school teacher may be defined as a set of related cognitions maintained for that position by teachers themselves and related others. Educators and others have often assumed that the teacher role consists of a consensual configuration of cognitive elements. In other words, it is assumed that all people see the teacher position in the same way regardless of positional affiliation.

If this were the case, then any problem encountered by the teacher in the fulfillment of role requirements would be the result of personal inability to meet these requirements. Conflicts could not arise due to the influence of incompatible expectations and values held by various persons. Any problem of adjustment the teacher might have could legitimately be attributed to the individual teacher because everyone else would know both what was expected and what should be done.

However, if consensus is not the case and there are a significant number of cognitive disparities between and within positions related to the teacher role, the implications are very different. The problems encountered by the teacher might very well be the result of converging incompatible cognitions in areas crucial to the teacher role. That there are several positions having a significant relationship to the teacher position is indisputable. Teachers interact with pupils, parents, and school officials, while the public at large pays the bills of educational systems.

The members of each of these different social positions are the recipients of differential socialization. It would, therefore, be very surprising if all members of every position held the same values and expectations for the teacher role. Even if they did display consensus on cognitive content, would

---

* Received in the Editorial Office on September 26, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The research reported here was part of a larger program conducted at the University of Kansas City and at the University of Missouri pursuant to contracts from the Cooperative Research Program of the Office of Education, Department of Health, Education and Welfare. Invaluable aid and assistance in the preparation of this article was given by the Oklahoma State University Research Foundation.

the various positions be oriented toward this content in the same manner? For example, would representatives of all social positions expect teachers to behave firmly in the classroom, or is it not likely that some would expect, like, or dislike very firm, moderately firm, or even lenient behavior? A review of the literature on teacher role and role conflict reveals a paucity of information on such issues.

These questions and others provided the focus for an empirical study of teacher role conflict recently completed at the University of Kansas City (9). This study, part of a larger program of research on role theory and the role of the public school teacher, was conducted in an attempt to operationally determine the extent of cognitive disagreement among four social positions (teachers, parents, pupils, and school officials) on what teachers do, should, and should not do. Other reports of the larger project on the conceptual structure of role theory and the empirical study of teacher role are also available (1, 3, 4).

B. METHOD

As various researchers and theorists have pointed out, the behavior of others may be appraised in several ways (5, 6, 7). In role theory, two major approaches have been studied extensively, the descriptive and the prescriptive orientations. The descriptive orientation deals with the individual's assessment of reality, or his picture of things as he presumes they are. A descriptive cognition applied to the behavior of a person or position is termed an *expectation*. Expectations give no information relevant to values and may not provide information relevant to actual behaviors unless it is assumed the reported expectation is veridical. The prescriptive orientation deals not with the assessment of reality, but rather with the stating of "oughts," or rights and wrongs for reality. Such a value oriented cognition may be called a *norm*. Norms may be positive or negative.

It was assumed, in the study herein reported, that norms (both positive and negative) and expectations are basic aspects or components of the teacher role held both by teachers themselves and by members of other positions connected with the educational system. Furthermore, it was assumed that one form of role conflict could be defined as the existence of disparate cognitions held for teachers, with these disparities being examined on an inter-positional basis.

1. *Sample and Population*

The setting for the study was a large metropolitan area encompassing five counties in two states and 406 individual public schools. A sample of 50 schools was selected from the total population on a stratified-proportionate sampling basis. This sample of schools was representative of the population in terms of community socio-economic type, state, county, grade level, number of teachers and number of pupils. From this sample of 50 schools, a total of 107 teachers (elementary and secondary), 375 parents, 375 pupils (ranging from Grade 6 through high school) and 98 school officials (including principals, superintendents and school supervisors) were selected as potential respondents. These potential respondents were apportioned through the 50 sample schools in terms of the number of teachers, parents, pupils and school officials represented in each school and district. Actual selection of respondents was by random techniques (using school identification lists and a table of random numbers). The final sample, from which data were actually collected, included 98 teachers, 261 parents, 237 pupils, and 67 school officials. It should be noted that the final sample does not represent a demographic sample of "the public." Rather, it is assumed that the final sample is representative of the population of positions sampled, all of which are oriented in a particular way toward the school system.

2. *Procedure*

Data for the study being reported were collected on a group interview basis with an instrument designed to measure expectational and normative cognitions held for teacher behaviors. Thirty different teacher behaviors in a variety of background contexts were chosen for this instrument on the basis of empirical pilot study findings. In the pilot study, respondents (similar in nature to the final sample of respondents) were asked (on open-ended instruments) to give lists of physical settings appropriate for teacher entrance, social situations likely to occur within these settings, and likely teacher behaviors within these settings and situations. Pilot study responses were analyzed and thirty behaviors were identified as being salient. These behaviors represented a variety of content and context and were felt to be fruitful for generating disparities between groups of respondents representing various social positions. (For other phases of the larger study dealing with the role of the public school teacher, fifty rather than thirty items were included in this instrument.)

These behavioral items were then matched with a rank-ordered, five-point scale of possible "degree of teacher behavior" alternatives. Each scale ran from the teacher doing "a great deal of," to "a lot of," "a moderate amount of," "a slight amount of," and "little or none" of the behavior involved.

A list of the final thirty items follows:

1. Watching pupils during study period
2. Supervising pupils in a classroom test
3. Singing during a classroom singing period
4. Ignoring pupils misbehaving in a classroom
5. Keeping pupils quiet in the library
6. Speaking out at PTA meeting in the auditorium
7. Policing the halls
8. Disciplining pupils during basketball game in the gym
9. Visiting with other teachers on the playground
10. Cooperating with other teachers in school
11. Visiting with other teachers on the playground
12. Watching for cheaters during a classroom test
13. Having pupils sing solos during a classroom singing period
14. Telling misbehaving pupils in the classroom to stop talking
15. Doing own studying and grading in the library
16. Attending PTA meetings in the auditorium
17. Allowing pupils to explore during experiments in the laboratory
18. Kibitzing during basketball game in gym
19. Supervising on the playground
20. Cooperating with the principal in school
21. Reading own books during a study period
22. Leaving the room during a classroom test
23. Grading papers during a classroom singing period
24. Threatening pupils misbehaving in the classroom
25. Choosing books for pupils in the classroom
26. Visiting with parents during PTA meeting in the auditorium
27. Being partisan during football game on athletic field
28. Doing menial tasks during basketball game in gym
29. Disciplining pupils in cafeteria
30. Participating in community activities out of school

To avoid respondent fatigue, these behavioral items were broken down into three sets. Thus, a given respondent answered questions dealing with only ten items and only one-third of all respondents answered any given question. Careful sampling was used, however, to insure random assignment of respondents to the three sets of questions.

The final instrument was used to elicit expectations, positive norms and negative norms for "teachers in general." Two sample questions from the study are presented below. The first question asks for the respondent's expectation and the second question asks for positive and negative norms. Both of these questions are based on Item 1. Respondents were instructed to circle the number corresponding to the answer that most nearly represented their judgment.

WHAT DO YOU EXPECT OF TEACHERS?

Watching Pupils During Study Period

I expect that during a study period most teachers would do:	A great deal of watching pupils	5
	A lot of watching pupils	4
	A moderate amount of watching pupils	3
	A slight amount of watching pupils	2
	Little or no watching pupils	1

WHAT DO YOU MOST LIKE AND LEAST LIKE TEACHERS TO DO?

Watching Pupils During Study Period

		Most like	Least like
I most like and least like teachers during a study period to do:	A great deal of watching pupils	5	5
	A lot of watching pupils	4	4
	A moderate amount of watching pupils	3	3
	A slight amount of watching pupils	2	2
	Little or no watching pupils	1	1

For each item in the instrument, three major varieties of cognitive discrepancy between positions were analyzed (expectations versus expectations, positive norms versus positive norms, and negative norms versus negative norms). The statistic used for these analyses was the Mann-Whitney z_U, a statistic that tests the difference between two ordinal distributions and is comparable to the *t* test of difference between means. However, it does not assume interval scale data, nor is it sensitive to non-normality of the distributions involved (2, 8). Statistical computations were performed on a Burroughs Datatron.

C. FINDINGS

The findings reported in tables below consist of significant mean differences between the various positions being compared. These tables include items sharing a common content and a systematic pattern of differences and are organized to show positional comparisons and types of cognitive discrepancy. No findings are reported unless they are significant at $p < .05$ or better. In interpreting the mean differences, a positive sign indicates that the first listed subject position had a "higher" mean response than did the second. Thus, if teachers scored 3.450 for a selected cognition and behavioral item, while parents scored 2.000 on the same question, the mean difference reported would be 1.450.

Findings were significant considerably beyond chance expectation. Of 540 comparisons, 71 differences were significant at $.01 < p < .05$, 42 significant at $.001 < p < .01$, and 21 significant at $p < .001$. All statistical tests were two-tailed.

Table 1 displays significant inter-positional disparities for behavioral items dealing generally with "maintenance of order." Six different behavioral items are included in this table, with the majority of significant discrepancies occurring for positive norms. Major discrepancies for all cognitions are between parents versus pupils or teachers.

For positive norms, findings reported in this table indicate that parents more than teachers and pupils, approve of the maintenance of order. This trend holds, even for one item implying the lack of order maintenance. Teachers and pupils, more than parents, approve of "ignoring pupils misbehaving in the classroom." Other positive norm discrepancies indicate that parents also approve of order maintenance more than school officials ("policing the halls," and "disciplining pupils in the cafeteria"). One significant positive norm discrepancy was found when comparing teachers against school officials. School officials approve of more supervising on the playground than do teachers.

Findings for negative norms are consistent with those for positive norms. Pupils dislike maintenance of order more than parents, teachers and school officials. Teachers dislike this activity more than do parents. For expectations, parents are found to expect more maintenance of order by teachers than either teachers or pupils.

Table 2 displays significant results for items relating to "deviancy control." A consistent pattern of differences may be observed for the three behavioral items included in this table, with major discrepancies occurring between parents

TABLE 1
SIGNIFICANT INTER-POSITIONAL DIFFERENCES FOR MAINTENANCE OF ORDER

	Positive norms	Negative norms	Expectations		Positive norms	Negative norms	Expectations
Behavior Item, Watching pupils during study period.				*Behavior Item,* Ignoring pupils misbehaving in the classroom.			
Tch-Par	—	—	—	Tch-Par	.493**	—	—
Tch-Pup	—	—	—	Tch-Pup	—	—	—
Tch-S. Off	—	—	—	Tch-S. Off	—	—	—
Par-Pup	—	—.963***	—	Par-Pup	—.826**	.904***	—
Par-S. Off	—	—	—	Par-S. Off	—	—	—
Pup-S. Off	—	—	—	Pup-S. Off	—	—	—
Behavior Item, Keeping pupils quiet in the library.				*Behavior Item,* Policing the halls.			
Tch-Par	—.527**	—	—.417**	Tch-Par	—.911***	1.272***	—.673**
Tch-Pup	—	—.467*	—	Tch-Pup	—	—	—
Tch-S. Off	—	—	—	Tch-S. Off	—	—	—
Par-Pup	.504**	—.648***	.406**	Par-Pup	.731**	—1.561***	—
Par-S. Off	—	—	—	Par-S. Off	.641*	—	—
Pup-S. Off	—	.645*	—	Pup-S. Off	—	1.003*	—
Behavior Item, Supervising on the playground.				*Behavior Item,* Disciplining pupils in the cafeteria.			
Tch-Par	—.863***	.883***	—	Tch-Par	—1.103***	—	—.812*
Tch-Pup	—	—	—	Tch-Pup	—	—	—
Tch-S. Off	—.679*	—	—	Tch-S. Off	—	—	—
Par-Pup	.602**	—1.150***	—	Par-Pup	—	—	—
Par-S. Off	—	—.780**	—	Par-S. Off	.744*	—	—
Pup-S. Off	—	—	—	Pup-S. Off	—	—	—

NOTE: For all tables included in this paper, differences significant at $.01 < p < .05$ are indicated with a single asterisk, at $.001 < p < .01$ with two asterisks, and at $p < .001$, with three asterisks.

TABLE 2
SIGNIFICANT INTER-POSITIONAL DIFFERENCES FOR DEVIANCY CONTROL

	Positive norms	Negative norms	Expectations
Behavior Item, Watching for cheaters during a classroom test.			
Tch-Par	—	.938*	—
Tch-Pup	—.606*	1.173**	—
Tch-S. Off	—	—	—
Par-Pup	—	—	—
Par-S. Off	.764*	—	—
Pup-S. Off	1.048**	—	—
Behavior Item, Threatening pupils misbehaving in the classroom.			
Tch-Par	—.830*	1.442**	—
Tch-Pup	—	1.055*	—
Tch-S. Off	—	—	—
Par-Pup	—	—	—
Par-S. Off	1.380**	—	—
Pup-S. Off	1.032*	—	—
Behavior Item, Telling misbehaving pupils in the classroom to stop talking.			
Tch-Par	—	—	—
Tch-Pup	—	—	—.563**
Tch-S. Off	—	—	—
Par-Pup	—	—	—
Par-S. Off	.717*	—1.095*	.706*
Pup-S. Off	.701*	—	.974**

and pupils versus teachers and school officials. Findings for positive norms indicate that parents and pupils, more than teachers and school officials, approve of deviancy control. The negative norm findings are compatible and show that teachers, more than parents and pupils, dislike deviancy control. School officials, more than parents, disapprove of telling misbehaving pupils in the classroom to stop talking. Expectational comparisons indicate that parents and pupils expect to find more deviancy control by teachers than do teachers and school officials.

Findings relative to a series of behavioral items concerned with "teacher self-indulgence" are displayed in Table 3. Most of the discrepancies stem from disagreements between parents and pupils versus teachers and school officials. Positive norm comparisons indicate that parents and pupils, more than teachers and school officials, approve of teacher self-indulgence. Pupils, more than parents, approve also of teacher self-indulgence. Compatible findings may be observed for negative norms, where teachers, and to some extent, school officials, dislike teacher self-indulgence more than do parents and pupils. Parents and pupils also expect more self-indulgence than do teachers and school officials.

Table 4 includes significant inter-positional disparities for "community participation." Major discrepancies stem from disagreements between teachers versus parents, pupils, and school officials. Parents, pupils, and school officials, for positive norms, approve of more community participation by teachers than do teachers. And pupils approve of this activity for teachers more than do parents. For negative norms, the findings are consistent with this pattern. Teachers, more than parents and pupils, disapprove of community participation. Parents disapprove of community participation more than do pupils. Expectational findings indicate that parents, pupils, and school officials expect more community participation by teachers than do teachers. And parents expect more of this activity than do school officials.

Results for two behavioral items concerned with "teacher cooperation" are included in Table 5. Teachers approve of more cooperation than do pupils. The greatest number of significant discrepancies occur for negative norms, where it may be observed that pupils, more than teachers, parents and school officials, dislike teacher cooperation. For expectations, it may be seen that teachers, in contrast to pupils and school officials, expect more teacher cooperation.

Table 6 is composed of five behavioral items which share neither a common pattern of difference nor a commonality of content. Differences for each of these items are somewhat unique. For the item, "supervising pupils in

TABLE 3
SIGNIFICANT INTER-POSITIONAL DIFFERENCES FOR TEACHER SELF-INDULGENCE

	Positive norms	Negative norms	Expectations
Behavior Item, Reading own books during study period.			
Tch-Par	—.770*	—	—1.127**
Tch-Pup	—1.600***	1.450*	—1.507***
Tch-S. Off	—	—	—
Par-Pup	—	—.628*	—
Par-S. Off	.738*	—1.310*	.947**
Pup-S. Off	—	—	.936**
Behavior Item, Grading papers during a classroom singing period.			
Tch-Par	—	—	—.873*
Tch-Pup	—.948**	1.340*	—1.127*
Tch-S. Off	—	—	—
Par-Pup	—	—	—
Par-S. Off	—	—	—
Pup-S. Off	1.081*	—	1.210*
Behavior Item, Visiting with other teachers on the playground.			
Tch-Par	—.481*	—	—.467**
Tch-Pup	—	1.164**	—
Tch-S. Off	—	—	—
Par-Pup	—	.872**	—
Par-S. Off	—	—	—
Pup-S. Off	—	—1.595**	—
Behavior Item, Leaving the room during a classroom test.			
Tch-Par	—	—	—
Tch-Pup	—.731*	—	—.461*
Tch-S. Off	—	—	—
Par-Pup	—.495*	—	—
Par-S. Off	—	—	—
Pup-S. Off	—	—	—
Behavior Item, Doing own studying and grading in the library.			
Tch-Par	—	1.072*	—.625*
Tch-Pup	—	—	—.614*
Tch-S. Off	—	—	—
Par-Pup	—	—.628*	—
Par-S. Off	.738*	—1.310*	.947**
Pup-S. Off	—	—	.936**
Behavior Item, Visiting with other teachers during study period.			
Tch-Par	—	—	—
Tch-Pup	—.504*	—	—.667***
Tch-S. Off	—	—	—
Par-Pup	—	—	—
Par-S. Off	.577*	—	—
Pup-S. Off	.600*	—	—
Behavior Item, Singing during a classroom singing period.			
Tch-Par	—	.769*	—
Tch-Pup	—	—	—
Tch-S. Off	—	—	—
Par-Pup	—	—.724**	—
Par-S. Off	—	—	—
Pup-S. Off	—	—	—

TABLE 4
SIGNIFICANT INTER-POSITIONAL DIFFERENCES FOR COMMUNITY PARTICIPATION

	Positive norms	Negative norms	Expectations		Positive norms	Negative norms	Expectations
Behavior Item, Speaking out at PTA meeting in the auditorium.				*Behavior Item*, Attending PTA meetings in the auditorium.			
Tch-Par	—.798***	1.376***	—1.263***	Tch-Par	—.480*	—	—
Tch-Pup	—.732**	1.086*	—1.128***	Tch-Pup	—.859***	—	—
Tch-S. Off	—	—	—.739**	Tch-S. Off	—.856**	—	—
Par-Pup	—	—	—	Par-Pup	—.379**	—.241*	—
Par-S. Off	—	—	.524*	Par-S. Off	—	—	—
Pup-S. Off	—	—	—	Pup-S. Off	—	—	—
Behavior Item, Visiting with parents during PTA meetings in the auditorium.				*Behavior Item*, Participating in community activities out of school.			
Tch-Par	—	1.223*	—	Tch-Par	—	1.126*	—
Tch-Pup	—	—	—	Tch-Pup	—	—	—
Tch-S. Off	—1.342*	1.857*	—	Tch-S. Off	—	—	—
Par-Pup	—	—	—.674*	Par-Pup	—.778**	—	—.550*
Par-S. Off	—	—	—	Par-S. Off	—	—	—
Pup-S. Off	—	—	—	Pup-S. Off	—	—	—

TABLE 5

SIGNIFICANT INTER-POSITIONAL DIFFERENCES FOR TEACHER COOPERATION

	Positive norms	Negative norms	Expectations		Positive norms	Negative norms	Expectations
Behavior Item, Cooperating with other teachers in school.				*Behavior Item,* Cooperating with the principal in school.			
Tch-Par	—	—	—	Tch-Par	—	—	—
Tch-Pup	—	—.339*	—	Tch-Pup	.558**	—.672*	.387**
Tch-S. Off	—	—	—	Tch-S. Off	—	—	.278*
Par-Pup	—	—.271*	—	Par-Pup	—	—.990***	—
Par-S. Off	—	—	—	Par-S. Off	—	—	—
Pup-S. Off	—	—	—	Pup-S. Off	—	.898*	—

TABLE 6

SIGNIFICANT INTER-POSITIONAL DIFFERENCES FOR MISCELLANEOUS BEHAVIOR ITEMS

	Positive norms	Negative norms	Expectations
Behavior Item, Supervising pupils in a classroom test.			
Tch-Par	—	—	—
Tch-Pup	—	—.984**	.688*
Tch-S. Off	—	—	—
Par-Pup	.613*	—.926***	.688**
Par-S. Off	—	—	—
Pup-S. Off	—	1.149**	—1.155**
Behavior Item, Having pupils sing solos during a classroom singing period.			
Tch-Par	—	—	—
Tch-Pup	—	.814*	—
Tch-S. Off	—	—	—
Par-Pup	—	—	.337*
Par-S. Off	—	—	—
Pup-S. Off	—	—	—
Behavior Item, Doing menial tasks during basketball game in the gym.			
Tch-Par	—	.931*	—
Tch-Pup	—.912*	1.568**	—
Tch-S. Off	—	.429*	—
Par-Pup	—	—	—
Par-S. Off	—	—	—
Pup-S. Off	—	—	—
Behavior Item, Disciplining pupils, during basketball game in the gym.			
Tch-Par	—.576*	1.133**	—.576*
Tch-Pup	—.851**	—	—.917***
Tch-S. Off	—	—	—
Par-Pup	—	—.760*	—
Par-S. Off	—	—	—
Pup-S. Off	.620*	—	—
Behavior Item, Allowing pupils to explore during experiments in the laboratory.			
Tch-Par	—	.675*	—
Tch-Pup	—.696**	—	—
Tch-S. Off	—.682*	—	—.744*
Par-Pup	—.533**	—	—
Par-S. Off	—	—	—
Pup-S. Off	—	—	—

a classroom test," parents show more approval than pupils. Other findings for this item indicate that pupils dislike more and do not expect this activity as much as do teachers, parents and school officials.

For the item, "disciplining pupils during basketball games in the gym," parents and pupils approve and expect to find this activity more than do teachers. Pupils approve of this behavior more than do school officials. Teachers and pupils exhibit greater disapproval of this activity than do parents.

For the behavioral item, "having pupils sing solos during a classroom singing period," teachers show greater disapproval than pupils and parents and expect to find this activity more frequently than do pupils.

Findings for the item, "allowing pupils to explore during experiments in the laboratory," indicate that pupils and school officials exhibit more approval than teachers. Also, pupils exhibit more approval than parents. For negative norms, it may be seen that teachers disapprove of this activity more than do parents. School officials expect more of this particular behavior on the part of teachers than do teachers themselves.

For the item, "doing menial tasks during basketball games in the gym," major disagreements are between teachers and all other positions. Pupils approve of the performance of menial tasks more than do teachers. Teachers, as shown by negative norm comparisons, disapprove of menial task performance more than do parents, pupils and school officials. No significant discrepancies are found in expectations for this behavioral item.

D. Summary and Discussion

Data reported in this paper have been based on a comparison of cognitions held for selected teacher behaviors by members of four school-related social positions. It seems clear that a number of significant disparities exist among teachers, school officials, parents, and pupils over teacher role cognitions. These disparities were found for positive norms, negative norms and for expectations, and occurred in a variety of content areas.

Among the four positions examined, pupils were found to be the most dissident, with teachers and school officials standing closest together in terms of role elements held. Parents tended to fall between school personnel and pupils. In general, most of the disagreements among positions centered on behaviors having to do with teacher self-indulgence, maintenance of order, and community participation.

The study herein reported was initiated under the assumption that interpositional disparities constituted "role conflict for the teacher" in the sense that the teacher was caught between conflicting pictures or values for her behavior. It is useful to examine this assumption.

It seems unlikely that the disparities between pupil and adult role cognitions on the thirty items of this investigation present a salient problem for teachers. If age is the major factor responsible, pupils may be assumed to "grow out of it" in time. In addition, despite their prevalence in the school, pupils do not have much real power in the determination of school events. The pupil who is faced with a teacher who violates his norms or expectations is ordinarily powerless to take action against the individual teacher or against the school system.

The reverse is true for parent-teacher, parent-school official, or school official-teacher disparities. In each case, the actors have considerable power and may, under certain circumstances, exercise it. The parent, for instance, who is incensed at the activities of an elementary teacher may create all sorts of trouble for that teacher or for the school administrator who supports her. It is strongly implied, therefore, that such widespread disparities as are revealed in this study are a potential source of irritation for public school teachers.

At the same time, it is not clear from these data that teachers adequately "perceive" these conflicting disparities. It is possible, as an example, for teachers to be unaware of existent disparities, suggesting that deviant teachers might receive the spleen of offended school officials, parents or other teachers without an understanding of the cause for annoyance. Similarly, some teachers may attribute conflicting norms or expectations where no such conflict exists. The authors are currently engaged in additional research designed to explore these latter issues more thoroughly.

But role conflict for the teacher has additional implications for the school official, the school system and the wider society. Many of the findings reported here suggest that parents are simply uninformed of much that goes on in the classroom or that they pick up misinformation from their children. When mobilized, these misinformations can prove a threat or harassment to the official or school board member and can disrupt the serious work of the school. Or, when teachers and school officials differ significantly on standards for teacher behavior, conflicts can appear within the fabric of the institution itself.

Perhaps more seriously, with severe role conflict no objective evaluation of teacher performance is possible. When parents, teachers, and school officials differ in standards of judgment for teacher behavior, then one judger's effective teacher is equivalent to another judger's anathema.

The incidence of wide-spread disagreement on the role of the teacher suggests that the school is undergoing rapid change or that persistent incongruencies may exist in the school as a stable social institution. Whichever interpretation is favored, it is implied that teachers, parents, pupils, and school officials alike must suffer in considerable part from lack of agreement over such issues as maintenance of order, deviancy control, and community participation by teachers.

REFERENCES

1. BIDDLE, B. J., ROSENCRANZ, H. A., & RANKIN, E. F., JR. Studies in the Role of the Public School Teacher. Vol B: Bibliographies on Role Terms, Role Conflict, and the Role of the Teacher, I: Orientation, Methods and Materials, II: General Characteristics of the School Teacher's Role, III: Positional Differences in Teacher Role, IV: Role of the Teacher and Occupational Choice, V: Own and Attributed Cognitions for the Teacher. Columbia: Univ. Missouri Press, 1961.
2. BIDDLE, B. J., & SIMPSON, A. Studies in the Role of the Public School Teacher. Vol. C: A Program for the Processing of Ordinal Data and Computation of Significance for Selected Central Tendency Differences. Columbia: Univ. Missouri Press, 1961.
3. BIDDLE, B. J., & TWYMAN, J. P. Concepts of Role and Role Conflict. Unpublished manuscript, 1962. (Mimeographed.)
4. BIDDLE, B. J., TWYMAN, J. P., & RANKIN, E. F., JR. The Role of the teacher and occupational choice. *School Review,* 1962, **70**, 191-206.
5. GROSS, N., MASON, W. A., & MCEACHERN, A. W. Explorations in Role Analysis. New York: Wiley, 1958.
6. PARSONS, T. The Social System. Glencoe, Ill.: Free Press, 1951.
7. ROMMETVEIT, R. Social Norms and Roles. Minneapolis: Univ. Minnesota Press, 1954.
8. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
9. TWYMAN, J. P. Role Conflict in the Teaching Profession. Unpublished doctoral dissertation. The University of Kansas City, 1961.

Department of Education
Oklahoma State University
Stillwater, Oklahoma

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 199-209.

A NOTE ON "STRENGTH OF ASSOCIATION"*

Swarthmore College[1]

Solomon E. Asch and Marged Lindner

A. Purpose

The concept "strength of association" has a place in the psychology of human learning second only to that of association itself. It has been an axiom that phenomena of acquisition and loss refer to changes of associative strength. Since the goal of experimentation in this region requires as a rule the measurement and comparison of levels of acquisition and recall, there is hardly an investigation that fails to presuppose this property of associations. We propose to show that the automatic application of the concept of associative strength contains a serious error, that it neglects an alternative interpretation of the effects of practice and forgetting. An examination of the source of this error raises a more far-reaching question, namely, whether associations once formed increase at all in strength with further practice.

Basic to the present discussion is a recent analysis of the relation of association to recall by Asch and Ebenholtz (1), the principal conclusion of which is that associative recall is a function (among others) of two distinct processes, of recall as such and of the associative connection. Let us illustrate the analysis by referring to the problem to which it was first applied, namely, the directionality of associations.

It has been a frequently reported finding that associative recall in the forward direction (from *a* to *b*) occurs with higher frequency than in the backward direction (from *b* to *a*).[2] The accepted interpretation of this finding has been that forward and backward associations differ in strength; the justification for this interpretation is the axiom that frequency of occurrence of associative recall is an index of relative strengths of associations. The study referred to raised the following doubt concerning this formulation. The evidence in question consists of data of recall, but the inference about directionality refers to an event preceding recall, namely, the formation of associations. This indirectness would be of little consequence if recall merely

---

* Received in the Editorial Office on October 1, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a grant from the Ford Foundation. We wish to acknowledge the help of Mr. Keith Johnson with these experiments.

[2] We designate the first and second members of an association as *a* and *b*, and avoid describing them as "stimulus" and "response" for a reason that will be explained below. (See pp. 206-207.)

reflected the associative situation. This is indeed the assumption central to an associative theory of recall, but there is the following reason to question it. The study of associative directionality bases itself largely on the mastery of paired associates, employing the method of anticipation. Given these conditions, *S* is constantly anticipating, that is, recalling, the *b* terms during learning but not the *a* terms. The subsequent forward test, usually a test of aided recall, asks for the recall of the same *b* terms that had been previously anticipated; but the backward test requires the recall of the *a* terms that were not previously anticipated. One must, therefore, consider the possibility that the inequality of forward and backward performances is a function not of differences of associative strength but of differential recall or availability of the terms to be produced, or that the procedure of learning has systematically favored recall in one direction. If the obtained asymmetry is one of recall, not of association, it would follow that no conclusion can be drawn about the relative strength of two associations unless the availability of the terms to be produced is known. In accordance with this analysis, Asch and Ebenholtz compared forward and backward associative recall when the availability of the *a* and *b* terms was made approximately equal (by a procedure of prefamiliarization). Under these conditions, differences between forward and backward recall were either abolished or reduced to a small amount. This evidence was the basis for the formulation of a principle of associative symmetry, that is, of the equality of forward and backward association.

The analysis just described tells us that an operation of measurement hitherto interpreted as evidence for a difference of associative strength is a function of a different process, one of recall. It reveals the error of the unqualified proposition that, given two associations, the one that occurs with higher frequency is necessarily the stronger. In short, it throws doubt on the operational criterion of frequency of occurrence as an index of associative strength.

The investigation to be described now provides a further, and in one respect more forceful, demonstration for this conclusion. Briefly, we will obtain different levels of acquisition in two groups on an associative task, and we will show that the groups are nevertheless equivalent with respect to the associations they have formed.

B. EXPERIMENT

1. *Procedure*

The materials of learning were a set of eight paired associates. These consisted of eight one-place numbers (from two to nine), each paired with three letters. Four of the letter combinations were one-syllable words (CUT, MOP, WIN, YES); the other four were trigrams (HFG, KBV, RZQ, JAD), chosen from Underwood and Schulz (5, pp. 244-245). The frequency of occurrence of the trigram letter sequences in English was low; except for the last trigram the pronunciation values were also low. The words were of high frequency; according to the Thorndike-Lorge count (4), three of the words were of AA frequency. Thus, the letter sequences consisted of two sets differing in meaning, pronunciability and frequency of occurrence in English. To reduce intra-serial interference, no letter appeared more than once in the list.

There were two conditions of learning. For Group 1 the words and trigrams were the first members of each pair; for Group 2 these were the second members of each pair. In all other respects the conditions of learning were identical for both groups. Learning was by the method of anticipation. The first member appeared alone in the memory drum for three seconds, followed immediately by the pair for an additional three seconds; the inter-trial interval was nine seconds. *S* was instructed to pronounce the numbers and to spell the trigrams and words. The first member of the pair was recited aloud when it appeared alone; the second member was to be anticipated (after the first trial), or to be recited upon its exposure if *S* failed to anticipate. All *S*s received equal practice—one trial for the presentation of the pairs, followed by three anticipation trials.

There were two different pairing arrangements of syllables and numbers, one for each half of the *S*s in each group. Within each arrangement the pairs appeared in four different orders, and these were presented on successive trials. The sequence of the orders was systematically varied from *S* to *S*.

A test of aided recall followed immediately after the last learning trial. There were two recall conditions. One half of each group was tested in the forward direction, the direction in which they had anticipated during learning. The other half of each group was tested in the backward direction; they were shown the second member of each pair and were instructed to recall the corresponding first member. The respective items appeared in the memory drum at six second intervals; *S*s were encouraged to guess when uncertain.

The combined learning-recall conditions yielded four subgroups (see Table 1). *S*s were assigned to the respective subgroups in the order of their appearance in the laboratory. There was a total of forty *S*s, 10 in each subgroup, all Swarthmore College students.

C. RESULTS

Table 1 contains the learning and recall scores, separately for pairs containing trigrams and words. The index of learning was correct anticipations on the last (or fourth) learning trial. Recall scores were in terms of number of correct recalls. Partial recalls were few; their exclusion has a negligible effect on the analysis that follows.

TABLE 1
LEARNING AND RECALL SCORES

Group	N	Learning task	Total correct anticipations on last learning trial		Recall task	Total correct anticipations on recall test	
			Tri-gram pairs	Word pairs		Tri-gram pairs	Word pairs
1A	10	L-N	16	27	Forward: L-N	22	31
1B	10	L-N	17	27	Backward: N-L	10	24
2A	10	N-L	6	22	Forward: N-L	11	25
2B	10	N-L	6	26	Backward: L-N	17	31

L = letters; N = number. Group 1 associated letter-number pairs, Group 2 associated number-letter pairs. Groups 1A and 2A were tested for recall in the forward direction, Groups 1B and 2B in the backward direction.

1. *Learning*

An analysis of variance of the learning scores was done by means of a Lindquist Type I design (3), with two variables used in classification: *A*. Trigram vs. word, *B*. Number-letter vs. letter-number sequence. The variation contributed by *A* applies within *S*s, that contributed by *B* between *S*s. The analysis is summarized in Table 2.

TABLE 2
ANALYSIS OF VARIANCE OF LEARNING SCORES

Source	*df*	*MS*	*F*
Within *S*s			
Trigram vs. Word (A)	1	40.6	81.2***
Trigram vs. Word × Sequence during learning (A × B)	1	3.0	6.0*
Error (within)	38	0.5	
Between *S*s			
Sequence during learning (B)	1	9.1	11.4**
Error (between)	38	0.8	

* $P < .05$.
** $P < .01$.
*** $P < .001$.

The two learning conditions (Groups 1 and 2) differed significantly in level of acquisition. The difference was almost entirely restricted to pairs containing trigrams; correct anticipation of trigrams was only about one-third that of numbers (12 vs. 33). Correct anticipation of words was only slightly lower than of numbers (48 vs. 54). Pairs containing trigrams produced lower learning scores than pairs containing words, whether these were the first or second members. However, the difference was far larger when the trigrams were to be anticipated (12 vs. 48 for Group 2) than when they were the first members of their pairs (33 vs. 54 for Group 1).

2. *Recall*

An analysis of variance of the recall scores was performed, employing a Type III design (3). Three variables were used in classification: *A*. Trigram vs. word; *B*. Number-letter vs. letter-number sequence during learning; *C*. Forward vs. backward recall.

TABLE 3
ANALYSIS OF VARIANCE OF RECALL SCORES

Source	*df*	*MS*	*F*
Within *S*s			
Trigram vs. Word (A)	1	32.5	37.4***
$A \times C$	1	0.4	
$B \times C$	1	0.4	
$A \times B \times C$	1	0.2	
Error (within)	36	0.9	
Between *S*s			
Sequence during learning (B)	1	0.1	
Forward vs. Backward recall (C)	1	0.6	
$B \times C$	1	12.0	7.7**
Error (between)	36	1.6	

** $P < .05$.
*** $P < .001$.

Table 3 shows that the only significant sources of variation were *A*, or differences due to the greater difficulty of trigrams as compared with words, and $B \times C$, or the interaction between sequence during learning and direction of recall (forward vs. backward) in relation to learning. Neither the sequence during learning (*B*) nor the direction of recall (*C*) contributed significantly to the total variance. The interaction $B \times C$ corresponds of course to the sequence of number-letter vs. letter-number during recall. In short, the fact that trigrams were less easily recalled than words, and that the letter-number sequence represented a lower order of difficulty than the number-letter sequence, suffices to account for the observed differences. As we saw, the learning data showed the same effects.

D. Discussion

The point of this investigation may be briefly summarized. Two groups mastered two associative tasks; these involved the formation of associations between the same pairs of terms, but in opposed directions. The learning performances of the two groups differed markedly, but they performed nearly identically when the subsequent test of recall was identical.

1. *Principle of Associative Symmetry*

Before turning to the further implications of these findings, we note that they provide added support for the principle of associative symmetry. They also clarify a point unresolved by the study of Asch and Ebenholtz (1). The latter investigation proceeded in two steps. First, it confirmed the frequently reported finding that forward recall is superior to backward recall under the standard conditions of paired associate learning, when differential item availability obtains. Second, in order to decide whether the obtained asymmetry was a function of association or of recall, further experiments were performed in which the items were familiarized (by means of free recall learning) before paired associate learning. The latter procedure eliminated or strongly reduced the difference between forward and backward recall. This result demonstrated associative symmetry when item availability was equalized *before* the formation of associations; it did not strictly establish the equality of forward and backward association under conditions of unequal item availability at the time of learning. The present study tells us that the formation of associations functioned identically in two groups when there was, during learning, unequal availability for the terms to be associated, and when the relative availability of the first and second members of associated pairs was reversed. Indeed, we now relied upon differential item availability during learning (instead of minimizing it) in order to demonstrate associative symmetry. The findings thus extend the range of conditions under which associative symmetry obtains.

2. *Distinction Between Association and Recall*

At the center of the preceding analysis is the proposition that recall and association are separate processes. More specifically, we hold that recall is a function of availability, the latter term referring to a property of memory traces for data, and that the conditions of availability are not those of association. The proposition that there are operations of recall independent of association comes into direct opposition with the assumption of associative theories that recall occurs only via associative connections. In support of an availability theory of recall one may cite a recent study by Asch and Ebenholtz (2) that demonstrated the occurrence of consecutive free recall of data in the absence of associative connections between them. This study also identified some of the determinants of availability, among them serial recency and primacy, and frequency of prior recall. On this basis the writers concluded that there are non-associative conditions of free recall. The present investigation, and that of Asch and Ebenholtz, further establish that recall of data is to a marked degree distinct from associations also when the latter are present. These findings cannot be reconciled with theories that treat of recall solely in terms of association.

The principle of symmetry illustrates one consequence of the abandonment of an exclusively associative theory of recall. It should be clear that the problem of associative directionality does not admit of a solution without the separation between association and recall. Indeed, in the absence of this theoretical step the problem cannot be properly formulated. The consequence of the assimilation of recall to association has been to focus investigation on increasingly more careful comparisons between forward and backward performances, on the assumption that there is one correct relation to be discovered. In terms of the present analysis this purely empirical effort is misguided; given the independent character of recall, the relation between forward and backward recall must vary as a function of availability, at the same time that association in the two directions is equal. The equating of recall with association also forbids the adequate interpretation of empirical results. When applied to the present investigation, this position requires the strange conclusion that backward association can be weaker than forward association, but that it can also be stronger, and that in still other instances they are entirely equal (Table 1). The preceding analysis resolves these apparent contradictions.

3. *Consequences for Concept Strength of Association*

The separation of association and recall has the most direct consequences for the problem of associative strength. It tells us that differences in availability have been erroneously attributed to strength of association. In the present study two conditions of learning an associative task produced strong, consistent differences. The accepted interpretation of the finding that trigram-number pairs produced nearly three times as many correct anticipations as number-trigram pairs is that the former associations were easier to form than the latter. The further finding that this difference during learning nevertheless produced similar or identical levels of recall demonstrates the error of this interpretation, and compels the conclusion that the difference at the time of learning was one of performance, not of association.[3] These observations direct attention to the error of defining the presence and strength of an association in terms of the probability of occurrence of a response. An unqualified response-defined criterion of association slights the fact that recall has properties distinct from those of association, and that these may vary independently of each other.[4]

The preceding analysis raises a fundamental question about the concept "strength of association" itself. One must now ask whether differences in the availability of terms may not account *completely* for the effects hitherto attributed to associative strength, that is, whether an association once formed increases at all in strength. The possibility has to be faced that once an association has formed it exists at full strength, and that further practice exerts effects solely on availability. The assumption of strength of association is indeed problematical when one considers that the conditions of practice said to produce increasing strength of association involve as a rule increase of availability, and that the conditions from which loss of association is inferred generally entail loss of availability. To be sure, the present study does not provide an answer; changes of availability do not necessarily exclude changes of associative strength. The direction a solution must take is, however, clear. If one accepts the proposition that recall plays an independent part in associative reproduction, all conclusions about strength of association require that the contribution of recall proper be known. Investigation of the validity of the concept "strength of association" is now in progress.

4. *Consequences for Stimulus-Response Conceptualization of Associations*

Modern associationism has become part of behavior theory. It bases itself upon the phenomena and interpretations of conditioning, the principal terms of which are stimulus, response, and stimulus-response connection. The paired-associate experiment provides a clear illustration of the extension of these ideas to the problems of association. The first and second members of the pair are assigned, respectively, the status of stimulus and response; and a correct anticipation is said to be evidence of stimulus instigation.

---

[3] The conclusion that inferences about the ease or strength of association are not valid in the absence of knowledge concerning recall has a wide application. First, it is not limited to the index of frequency, upon which we have concentrated. Other measures that have been employed as indices of associative strength, such as latency, are also functions of recall, and the preceding analysis may hold for them. Second, the conclusion is relevant to all associated contents. Thus one may not maintain that there are associative differences between visual and auditory data, or between the data of other modalities, without knowledge of the role of recall.

[4] For a discussion of the operational definition of associations. (See **1**, pp. 138-39.)

One consequence of the present work is to rob the stimulus-response conceptualization of relevance for the theory of associations. The application of these terms to the phenomena of association can be justified only on the ground that they have distinct properties in the associative process. The findings of Asch and Ebenholtz (1), and their confirmation in this study, demonstrate the functional equivalence of the terms entering into association. Once it is shown that a "response" can produce a "stimulus" as well as a "stimulus" can produce a "response," the basis for this distinction disappears. One's faith in this conceptualization is hardly enhanced by the further finding, reported above, that backward recall can be higher than forward recall, or that the "response" can apparently be superior to the "stimulus" in eliciting its associate.

We conclude that these categories are without relevance to the phenomena of association. To be sure, the respective terms of an association can acquire distinct functional properties, but one major difference between them is, we saw, that of availability, which refers to a property of the memory traces of items, not of associations.

In the light of these conclusions it becomes necessary to reexamine the concepts in the area of association and memory, and the interpretations deriving from them. As a further illustration of the consequences of the present analysis, we cite one aspect of the recent work of Underwood and Schulz (5). These writers have proposed a two-stage process of associative acquisition, the first of which consists of making the "response" terms available for association, and the second of the formation of the associative connection. Relevant to the present discussion is the assumption that an item must be available, that is, recallable, in order to enter into the associative connection, or that "response availability" determines "the order in which units start entering into the second or associative stage of learning" (5, p. 286). The findings reported in this paper are in contradiction with this proposition. The equivalence of forward and backward association under conditions of unequal availability establishes beyond question that associations are formed *before* the relevant term can be produced. The well-known superiority of the method of matching as a test of association to the method of aided recall is equally decisive evidence for this point. At issue is precisely the distinction that has concerned us in this discussion, that between association and recall. Failure of recall is not identical with the absence of associations, any more than the presence of recall is identical with the presence of associations.

E. SUMMARY

An associative task was constructed of pairs of terms differing in ease of recall. Two groups learned the task by the method of anticipation, one with the more difficult items, the other with the easier items as the second members of the pairs. There was a marked difference of learning performance between the two groups.

One half of each group was subsequently tested in the forward direction, the other half in the backward direction. Given the identical items as a test of recall, there was no significant difference between those who recalled in the forward and backward directions. This finding confirms and extends the evidence reported earlier by Asch and Ebenholtz in support of the principle of associative symmetry, or of the equality of forward and backward association.

The evidence for associative symmetry requires the assumption that associative recall is a function of two independent processes, one of recall and the other of association. The problem of associative directionality can neither be solved nor properly stated without the separation of recall and association. Contemporary theory, which defines strength of association in terms of the probability of occurrence of a response on an associative task, systematically obscures this difference.

The accepted interpretation of a difference in learning performance between two groups on an associative task is that they differ in strength of association. The finding that groups differing in the "learning" of an associative task did not differ in recall disproves the generality of this inference. It demonstrates that effects hitherto attributed to strength of association are a function of differential item availability.

The findings raise the further question, now under investigation, whether associations once formed can increase in strength with further practice.

The symmetry of associations removes the ground for the conceptualization of an association as a stimulus-response connection. The equivalence of both terms in evoking an association, and the evidence that backward recall may under some conditions be superior to forward recall, cannot be reconciled with the stimulus-response account.

The findings here reported require a reexamination of the body of concepts and procedures based upon the stimulus-response analogy.

REFERENCES

1. ASCH, S. E., & EPENHOLTZ, S. M. The principle of associative symmetry. *Proceed. Amer. Philos. Soc.*, 1962, **106**, 135-163.
2. ———. The process of free recall: Evidence for non-associative factors in acquisition and retention. *J. of Psychol.*, 1962, **54**, 3-31.
3. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. Boston: Houghton Mifflin, 1956.
4. THORNDIKE, E. L., & LORGE, I. The Teacher's Word Book of 30,000 Words. New York: Teachers College, Columbia Univ., 1944.
5. UNDERWOOD, B. J., & SCHULZ, R. W. Meaningfulness and Verbal Learning. Philadelphia: Lippincott, 1960.

Department of Psychology & Education
Swarthmore College
Swarthmore, Pennsylvania

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 211-221.

A MULTIVARIATE ANALYSIS OF DECISION MAKING AND RELATED MEASURES*[1]

VA Hospital, Perry Point, Maryland, and The Catholic University of America, Washington, D.C.

JAMES M. ANKER, JOHN C. TOWNSEND, AND JAMES P. O'CONNOR

A. PROBLEM

The rather scant literature on psychological studies of decision making may be due to the fact that the area still is in the formative stage, with more effort being expended upon general theory and classifying and making observations more reliable, than upon interrelating variables in the usual experimental format (4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 20, 22, 30). Several papers have suggested that probability preference is related to achievement motivation (1, 2, 21). Block and Petersen (3) elaborated the personality characteristics of subjects who behaved with obvious confidence, caution, and/or speed in a decision making situation involving a psychophysical judgment. Scodel, Ratoosh, and Minas (24) studied several psychological correlates of risk taking behavior. They state, "In general, our total results, though far from overwhelming, do point to the importance of personality variables in risk-taking. We conclude not only that theories of risk-taking that ignore such variables are grossly inadequate as models for prediction but that many personality correlates of risk-taking behavior remain to be discovered (24, p. 27)."

This paper reports the results of an exploratory factor analysis of a 49 × 49 intercorrelation matrix composed of carefully selected measures of decision making parameters, intelligence, personality, work values, and interest categories. The purpose of this study was threefold: (*a*) to reduce a relatively large number of hypothetically related measures into sources of common variance; (*b*) to select for closer inspection those factors upon which decision making measures loaded to determine mutual interrelationships with other measures in the matrix; and (*c*) to use these general factor

---

* Received in the Editorial Office on October 5, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study is one of a series in a research program in the area of decision making and related processes being pursued at The Catholic University of America under the auspices of ACF Industries, Riverdale, Maryland. The authors wish to express their gratitude to ACF Industries, particularly to J. R. Huffner, for their financial support, continued interest, and patience with this extensive research program.

descriptions as the origin of explicit hypotheses which might be investigated more profitably in the future by conventional experimental procedures.

B. PROCEDURE

Two-hundred-two male freshmen and sophomore Air Force ROTC students at The Catholic University of America were administered the following tests: Minnesota Multiphasic Personality Inventory (17); School and College Ability Tests (8); Work Values Inventory (23, 26); Strong Vocational Interest Blanks (25); and the Decision Making Ability Test.[2] Age ranged from 17 to 22 years with a mean of 18.6 years and a standard deviation of .81. The measures investigated are presented by factor grouping in Tables 2–7. The use of the MMPI as an objective personality measure, the SCAT as a measure of intelligence, and the SVIB as a measure of vocational interest patterns is well documented in the literature and no further elaboration of these measures will be given here. The Work Values Inventory and the Decision Making Ability Test, however, do merit further comment.

O'Connor and Kinnane (23) modified the Work Values Inventory developed originally by Super (26) in the Career Pattern Study by transforming the forced choice items into individual items to be rated on a four-point scale. The 30 items purportedly measure 15 work values or sources of satisfaction arising from work: social, theory, art, mastery, economic-material, creativity, planning-supervision, variety, independence, conditions of work, associates, boss, security, prestige, and way of life. O'Connor and Kinnane factor analysed these items and identified six values factors. No second order factor was found.

The Decision Making Ability Test was developed to provide a controlled "decision under risk" game situation in which various parameters of decision making might be measured: i.e., decision time; amount bet; accumulated amount won (or lost); frequency of wins; confidence of the subject in his decision; choice of advantageous, fair, or disadvantageous alternatives; and probability preferences. On each of 27 betting trials *S* was presented with a 4 × 6 matrix, the ordinate of which was made up of 4 amounts to be bet (1, 2, 5, or 10 points) and the abscissa of 6 different probabilities of winning (1:3, 1:4, 1:7, 1:14, 1:25, and 1:50). The matrix cell entries indicated the actual payoff on winning for any bet-odds combination chosen. Table 1 presents an illustrative matrix. Whether *S* won or not on any particular bet was randomly programmed within the apparatus to match fairly the odds chosen. A different bet matrix was used for each of the first 16 trials; one-third of the pay-offs greater than zero were advantageous, one-third fair, and one-third disadvantageous. Trials 17 to 27 were run with the same bet matrix, one in which all payoffs were fair. The decision making measures are reported separately for Trials 1–16 and Trials 17–27. The time from the initial presentation of the bet matrix to the final decision was recorded automatically in the apparatus. Before each bet was "played" *S* recorded his confidence in the bet winning on a percentile scale. Initially *E* described the test situation and the decision required, defined and illustrated advantageous, fair and disadvantageous bets, and presented a sample trial before tests trials began.

---

[2] Townsend, J. C., & Huffner, J. R. An apparatus for measuring decision making under risk. Paper in preparation.

The measures used were chosen because of their suspected relationships with decision making, as suggested by the literature and careful consideration of the decision making process. Many of the raw data distributions for the 49 measures appeared markedly skewed. While it was desired to use Pearson Product-Moment correlations in the subsequent analysis because of their power, the skewedness made this questionable. The use of computers[3] made it possible to do the analysis based on tetrachoric correlations (distributions dichotomized at the median) *and* Pearson Product-Moment correlations, and to compare the results. Although tetrachoric correlations lose a considerable amount of information and result in increased variability in the calculated statistic, they do normalize the distributions involved. Two intercorrelation matrices thus were generated for the 49 measures. Each was factor analysed by the complete centroid method (27) and the factor matrices analytically rotated to orthogonal simple structure using the normalized varimax solution (18).

C. RESULTS, DISCUSSION AND CONCLUSIONS

The factor analysis of the Pearson intercorrelation matrix produced 12 factors, that of the tetrachoric intercorrelation matrix 11 factors. Nine comparable factors were identified clearly in both solutions. In each case the remaining factors extracted were poorly defined and had low and/or multiple loadings. As expected, use of the tetrachoric correlations produced greater variability generally. They did not, however, systematically change the factor structure presented by the analysis of the Pearson correlations. The only obvious effect of using tetrachoric correlations in this case was to increase error variance of the correlations. For this reason the authors chose to use loadings based upon Pearson correlations to interpret the factors extracted.[4] Five factors included significant loadings (≥ ± .30) on decision making measures, the other four being factors that grew out of the interrelations between the other measures, and not directly related to decision making.

---

[3] The extensive computations in this study were done on an IBM 650 Computer using card programs for the several statistical analyses. The program for calculation of the intercorrelation matrix was developed by J. D. Hall, University of Indiana, Bloomington, Indiana. The complete centroid factor analytic program was written by Silvio O. Navarro, formerly at the Data Processing Center, A. and M. College of Texas, and currently the Assistant Director, Computing Center, University of Kentucky. The IBM card program for the normalized varimax rotation was a version received from the Computation Center, The University of Texas, Austin 12, Texas. The original machine language version of the program was obtained from the University of Michigan. Additionally, the authors express their appreciation to Mark Rhyner and Archibald Freelove of ACF Industries, Riverdale, Maryland, who wrote intermediate programs and directed the majority of computations done in this analysis.

TABLE 1
ILLUSTRATIVE "BET" MATRIX (TRIAL 14)

	Payoffs						Points bet
	0	0	0	0	0	0	10
	135	15	120	25	30	120	5
	54	14	98	1	5	23	2
	6	1	99	2	13	24	1
Odds	A	B	C	D	E	F	
	1:14	1:4	1:50	1:3	1:7	1:25	

Inasmuch as the focus of this study was upon the attributes of decision making and those personality variables associated with them, the four factors extracted in the analysis not containing decision making variables were only of incidental interest. Factor I, which we have called Adjustment-Maladjustment, much resembles the "General Maladjustment Factor" identified by Tyler (28) and the "Ego-Weakness Factor" identified by Kassebaum, Couch, and Slater (19). It accounted for 21.7 per cent of the estimated total common variance. Factors III, V, and VIII appear to be meaningful descriptions of vocational interest and value patterns which are consistent with the impressions of psychologists working in this field and using these instruments. They accounted for 13.8 per cent, 8.5 per cent and 4.6 per cent of the estimated total common variance respectively. The primary impact of these factors for this study is that they did not relate in any systematic way to the parameters of decision making investigated. These factors are presented in Table 2.

Although the validity of the MMPI scales legitimately has been questioned,

---

[4] The unrotated and rotated factor and transformation matrices are available from the senior author on request.

TABLE 2

Loading	Measure
	Factor I (Adjustment—Maladjustment)
.91	MMPI Pt scale
.90	MMPI Sc scale
.88	Manifest anxiety scale
.73	MMPI D scale
—.73	Ego strength scale
.72	MMPI Hs scale
.66	MMPI F scale
.60	MMPI Pa scale
.58	MMPI Pd scale
—.57	MMPI K scale
.54	MMPI Mf scale
.35	MMPI Hy scale
.35	MMPI Ma scale
	Factor III (Technical vs. Interpersonal Orientation)
—.94	SVIB technical occupations
.91	SVIB linguistics
—.77	SVIB masculinity-femininity
.72	SVIB occupational level
—.60	SVIB physical sciences
.47	MMPI Mf scale
—.44	SVIB business detail
.36	SVIB president-manufacturing concern
—.30	WVI security-economic-material
.30	WVI social-artistic
	Factor V (Cultural vs. Managerial Orientation)
.90	SVIB social welfare
.78	SVIB interest maturity
—.60	SVIB president-manufacturing concern
—.48	SVIB physical sciences
—.31	SVIB masculinity-femininity
	Factor VIII (Scientific Creativeness vs. Practical Orientation)
.70	SVIB biological sciences
—.58	SVIB business detail
.52	SVIB physical sciences
—.46	SVIB business contacts
.38	WVI heuristic-creative

as an objective personality measure it generally has fared better under close investigation than have others. Additionally, it is a widely used and accepted technique in spite of the problems that have been associated with it. Accepting the qualifications necessary for its use, it may be said that personality traits, i.e., traits associated with psychopathology, so measured show no consistent relationship with any of the parameters of decision making investigated in this study. The sample, however, consisted of college freshmen and sophomores in an Air Force ROTC unit and thus almost certainly would not show the variance in personality characteristics which might be found in the population. It would be reasonable to investigate this relationship in a sample of, or containing, diagnosed neurotic and psychotic subjects where such patterns would seem more likely to exist.

Factor II, decision time, as seen in Table 3, is composed only of measures of the time taken by *S*s to make a final decision after having been presented with the necessary information. *S*s were consistent in their performance between Trials 1–16 and Trials 17–27. The correlation of the sum and the median measures is, of course, spurious inasmuch as they came from the same data in each series of trials. This factor accounted for 10.7 per cent of the estimated total common variance. It is interesting, and rather surprising, to note that these measures are not consistently related to other decision making measures such as "goodness" of decision, decision confidence, and odds preference. Although there is some evidence to suggest a mild relationship between extremely fast decisions and poorness of decision (29), no over-all pattern was found here.

TABLE 3
FACTOR II
(Decision Time)

Loading	Measure
—.91	Sum of decision times for trials 1–16 minus 1,000
—.89	Median decision time for trials 1–16 multiplied by 10
—.83	Sum of decision times for trials 17–27
—.79	Median decision time for trials 17–27 multiplied by 10

High positive loadings on the total, verbal, and quantitative scores of the School and College Ability Tests clearly define Factor IV as an intelligence factor. These loadings are presented in Table 4. This is also the only factor upon which the "goodness" of decision measure loads. While this is only one of the attributes of decision making, it appears to be one of central importance. In this study it was defined by the *S*'s ability to discriminate between advantageous, fair, and disadvantageous bets. The factors extracted accounted for 33 per cent of its total variance. The odd-even reliability coefficient, corrected by the Spearman-Brown Formula, was .75. This suggests that approximately 42 per cent of the total variance of this measure is as yet not accounted for, but might be. That is to say that there are other variables systematically in-

TABLE 4
FACTOR IV
(Intelligence)

Loading	Measure
.96	SCAT total score
.70	SCAT verbal score
.66	SCAT quantitative score
.39	"Goodness" of decision

fluencing this measure that were not included in the original matrix of measures. It might be that a measure of abstract mathematical reasoning would account for more of this reliable variance than the one of relatively basic arithmetic skills found in the SCAT. Scales designed to describe variation in normal personality, as opposed to MMPI scales, also might account for some of this specific variance. These speculations await further investigation. The only variables measured in this study having a systematic relationship with "goodness" of decision were those of intelligence, and this relationship was clear. This factor accounted for 10.6 per cent of the estimated total common variance.

Factor VI is more difficult to interpret. The items loading on this factor are presented in Table 5. They suggest that the amount risked is a function of the immediately preceding "run of luck" and the *S*'s sensitivity to such events. Although the frequency of winning was randomly but fairly set to match whatever odds were chosen by *S*, irrespective of the odds some *S*s won more often than others. This phenomenon of "runs" obviously is true for any random series even though the over-all totals match chance expectancy. Sensitivity to such sequences of successes or failures apparently is reflected by the loading of the two WVI measures, social-artistic and security-economic-material. While this combination of values appears reasonable in this context, it is interesting to note that these two WVI measures generally have

TABLE 5
FACTOR VI
(Sequential Effects of Successes and Failures)

Loading	Measure
—.78	Points bet trials 1–16
—.74	Frequency of wins trials 1–16
—.45	WVI social-artistic
—.36	Points bet trials 17–27
—.35	Points won trials 1–16
—.35	WVI security-economic-material

a correlation of —.22. This factor seems to reflect the kind of behavior shown by a lover seeking a mate; bolder, riskier, more venturesome behavior following a series of successes, while some dimunition of ardor and inhibition of amorous behavior is likely to follow a series of failures. This factor, like the others, shows no systematic relationship with any other factor extracted. It accounted for 7.1 per cent of the estimated total common variance.

Factor VII, although defined by only two measures, appears clearly formed and readily interpretable. It is presented in Table 6. The two measures of decision confidence load highly on this factor and on no other. Although the reference axes were rotated orthogonally during the analysis, two dimensional sections were plotted to provide an estimate of the independence of the factors in hyperspace. With one exception these plots suggested that the orthogonal solution fit the factor space with unusual accuracy. The one exception noted was that there appeared to be a relationship between Factor IV and Factor VII. It would seem that intelligence and decision confidence are related mildly, and *negatively*. The estimated correlation would be in the order of —.20.

TABLE 6
FACTOR VII
(Decision Confidence)

Loading	Measure
.84	Mean confidence trials 17–27
.83	Mean confidence trials 1–16

As one might expect from previous research, a factor of odds preference was identified. Factor IX is presented in Table 7. This factor also appears genuinely independent of the others extracted. The items loading on this factor suggest that, other things being equal, choice of long or short odds is not related to other decision making factors, nor does it vary consistently with any of the personality traits, interests, or values measured. Inasmuch as achievement motivation already has been shown to be related to this factor

TABLE 7
FACTOR IX
(Odds Preference)

Loading	Measure
—.63	Frequency of wins trials 17–27
.56	Odds preference trials 17–27
.41	Points bet trials 17–27
—.34	Frequency of wins trials 1–16

it is interesting that none of the personality measures used in this study was related to it.

Five measures did not have significant loadings on any of the nine major factors identified: the number of points won or lost on Trials 17–27 of the decision making task; the Work Values Inventory factors associated with work conditions—associates, achievement—prestige, and independence—variety values; and the total number of A's and B's on the SVIB profile sheet.

Of the non-decision making variables included in the study, only the intelligence measures and two of the Work Values Inventory measures were found to be related to decision making. Although the scope of involvement is less than expected, this reduction of divers sources of data is unquestionably of value. Perhaps of greater significance is the finding that, with the one minor exception indicated above, the decision making factors identified are independent. The "goodness" of a decision within this sample is not related to the length of time it takes to make it, to *S*'s confidence in it, to preceding successes or failures, or to the odds preferred. In the application of these findings it appears that one could select decision makers on the basis of whichever one (or more) of the factors might be of value in a given situation. If decision speed is of crucial importance, for instance, it should be possible to choose decision makers on this basis without sacrificing "goodness" of decision. If sensitivity to preceding sequences of events is of importance it should be possible to select decision makers on this basis without producing decrements in the other decision making attributes.

One should be cautious, however, in generalizing these findings to conditions other than those under which this study was conducted. While the decision making task was developed carefully with special attention to the parameters of decision making of value to measure, there are a number of features which might limit its generality. *S*s were given a set number of points at the beginning of the task with which to gamble. Other than the loss of points there was no penalty for failure. How meaningful the "points" were is certain to have varied from *S* to *S* and there is no way of knowing the extent to which each *S* was involved. While *S*s generally appeared to be involved it is likely they viewed the task as an academic challenge with the point total reflecting their over-all success or failure. There is reason to believe the *S*s' responses might change under conditions involving substantial amounts of money or the possibility of damage to self or others following a poor decision. More generally, the influence of some general or global stress upon these types of decision making behavior remains to be investigated. Wins and losses in this study were fairly and randomly assigned on the basis of the odds chosen by *S* for each bet. The factor suggesting the importance of preceding series of successes or failures is rather difficult to interpret and certainly is open to question. In subsequent investigations it would be possible, of course, to vary systematically such sequences and evaluate effects on decision making factors. It also would be possible to vary interests and values experimentally for the same purpose, particularly in relation to different types of risk taking situations. Finally, the homogeneity of the sample restricts generalizations to comparable groups. Broader interpretation of these findings could be made only upon replication or extension of this study with representative samples.

D. Summary

Forty-nine measures of decision making, intelligence, personality traits, and vocational interests and values were administered to 202 college freshmen and sophomores. These measures were intercorrelated, factor analysed by the complete centroid method, and rotated analytically to orthogonal simple structure. Nine meaningful factors were identified, five of which contained decision making measures. These independent factors were identified as decision time, intelligence, sequential effects of successes and failures, decision confidence, and odds preference. The four other factors were identified as adjustment-maladjustment, technical vs. interpersonal orientation, cultural vs. managerial orientation, and scientific creativeness vs. practical orientation. The factors were discussed and their importance and implications for further research suggested. A number of cautions about the generality of the findings were presented, implying the need for further study.

References

1. Atkinson, J. W. Motivational determinants of risk-taking behavior. *Psychol. Rev.*, 1957, **64**, 359-372.
2. Atkinson, J. W., Bastian, J. R., Earl, R. W., & Litwin, G. H. The achievement motive, goal setting, and probability preferences. *J. Abnorm. & Soc. Psychol.*, 1960, **60**, 27-36.
3. Block, J., & Petersen, P. Some personality correlates of confidence, caution, and speed in a decision situation. *J. Abnorm. & Soc. Psychol.*, 1955, **51**, 34-41.
4. Cartwright, D. Decision-time in relation to differentiation of the phenomenal field. *Psychol. Rev.*, 1941, **48**, 425-442.
5. ————. The relation of decision-time to the categories of response. *Amer. J. Psychol.*, 1941, **54**, 176-196.
6. Cartwright, D., & Festinger, L. A quantitative theory of decision. *Psychol. Rev.*, 1943, **50**, 595-621.
7. Cohen, J., & Hansel, M. Risk and Gambling. New York: Philosophical Library, 1956.
8. Cooperative Test Division. School and College Ability Tests. Princeton: Educational Testing Service, 1957.

9. DALE, H. C. A. A study of subjective probability. *Brit. J. Statist. Psychol.*, 1960, **13**, 19-29.

10. EDWARDS, W. Probability-preferences in gambling. *Amer. J. Psychol.*, 1953, **66**, 349-364.

11. ———. The theory of decision making. *Psychol. Bull.*, 1954, **51**, 380-417.

12. ———. Probability-preferences among bets with differing expected values. *Amer. J. Psychol.*, 1954, **67**, 56-67.

13. ———. The reliability of probability-preferences. *Amer. J. Psychol.*, 1954, **67**, 68-95.

14. ———. Variance preferences in gambling. *Amer. J. Psychol.*, 1954, **67**, 441-452.

15. FESTINGER, L. Studies in decision: I. Decision time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. *J. Exp. Psychol.*, 1943, **32**, 291-306.

16. ———. Studies in decision: II. An empirical test of a quantitative theory of decision. *J. Exp. Psychol.*, 1943, **32**, 411-423.

17. HATHAWAY, S. R., & MCKINLEY, J. C. Minnesota Multiphasic Personality Inventory. New York: Psychological Corporation, 1943.

18. KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 1958, **23**, 187-200.

19. KASSEBAUM, G. G., COUCH, A. S., & SLATER, P. E. The factorial dimensions of the MMPI. *J. Consult. Psychol.*, 1959, **23**, 226-236.

20. LUCE, R. D., & RAIFFA, H. Games and Decisions. New York: Wiley, 1957.

21. MCCLELLAND, D. C. Risk taking in children with high and low need for achievement. *In* Atkinson, J. W. (*Ed.*), *Motives in Fantasy, Action, and Society.* Princeton, N.J.: Van Nostrand, 1958, 306-321.

22. MCGLOTHLIN, W. H. Stability of choices among uncertain alternatives. *Amer. J. Psychol.*, 1956, **69**, 604-615.

23. O'CONNOR, J. P., & KINNANE, J. F. A factor analysis of work values. *J. Couns. Psychol.*, 1961, **8**, 263-267.

24. SCODEL, A., RATOOSH, P., & MINAS, J. S. Some personality correlates of decision making under conditions of risk. *Behavioral Sci.*, 1959, **4**, 19-28.

25. STRONG, E. K. Strong Vocational Interest Blanks. Palo Alto, California: Consulting Psychologists Press, Inc., 1959.

26. SUPER, D. E. The Psychology of Careers. New York: Harper, 1957.

27. THURSTONE, L. L. Multiple-Factor Analysis. Chicago: Univ. Chicago Press, 1947.

28. TYLER, F. T. A factorial analysis of fifteen MMPI scales. *J. Consult. Psychol.*, 1951, **15**, 451-456.

29. VERDI, A. P. Decision making ability as a function of decision time, anxiety, and intelligence. Unpublished master's thesis, The Catholic University of America, 1961.

30. VON NEUMANN, J., & MORGENSTERN, O. Theory of Games and Economic Behavior. (Second ed.) Princeton: Princeton Univ. Press, 1947.

VA Hospital
Perry Point
Maryland

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 223-226.

DIMENSIONAL ANALYSIS IN APPLIED PSYCHOLOGICAL RESEARCH*

Lehigh University

P. JOSEPH LEHMAN AND EUGENE A. CRAIG

The present paper will describe a technique for the simplification and generalization of experimental procedure and for the efficient analysis of data that has demonstrated its value in engineering and that has important potential for applied psychology. The technique, known as dimensional analysis, may be traced back to Galileo but the assumptions and proof of the general theory were most clearly formulated by Buckingham in 1915 (1). Since that time it has become recognized as the most powerful way in which tests in fluid mechanics and heat transfer can be made compact in operating plan without loss in generality or control. A rigorous mathematical development of the theory may be found in Langhaar's 1951 book (2).

To apply this method, the investigator must know the number and kind of fundamental variables in his test. A fundamental variable is defined as any experimental variable that influences the test and can be changed independently of the other test variables. Once the fundamental variables have been identified, the experimenter can immediately reduce the number with which he must deal through the application of the Buckingham Theorem which states: "If any equation is dimensionally homogeneous, it can be reduced to a relationship among a complete set of dimensionless products" (2, p. 18). A dimensionally homogeneous equation is one whose form does not depend on the fundamental units of measurement. The dimensionless products are made up of the variables in such a manner that the dimensions cancel in each group.

This approach has many potential uses particularly in relation to aspects of engineering psychology and other applied areas where the emphasis may be on answering practical questions about a situation rather than on revealing its inner mechanism. The procedure can be illustrated by a problem in which the dimensions of the most efficient (minimum number of reading errors) circular dial gauge are desired. Consider a dial with the diameter of dial face, d; stroke-width of markers, w; spacing of markers, s; and length of

---

* Received in the Editorial Office on October 5, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

markers, l. One of the advantages of dimensional analysis is that by varying any of the four fundamental variables the experimenter could secure data to enable him to plot a single curve which immediately indicate the best combination of variables for the most efficient gauge. In addition, the curve could be used to assess the efficiency of any dial having the same fundamental and controlled variables.

In testing the dial we are interested in reading errors versus the geometry of the gauge so the equation will be of the form: Reading Errors, $R_e = f(l, s, d \text{ and } w)$. This equation, however, is not dimensionally homogeneous so both sides should be made dimensionless to satisfy the Buckingham Theorem. To non-dimensionalize the left side it is necessary to divide R_e by a quantity of the same units, such as the total number of readings, R_t. The right side can be non-dimensionalized by combining l, s, d and w in such a way that the units cancel. For example:

$$\frac{ls}{dw}, \frac{ld}{sw}, \frac{lw}{sd}, \frac{sd}{lw}, \frac{sw}{ld}, \frac{wd}{sl}.$$

Any of these parameters could be used but, considering the first combination, the final equation will be of the form: $R_e/R_t = f(ls/dw)$.

Suppose we have selected a dial of the following characteristics: $l = 0.5$ inches; $d = 12.0$ inches; $s = 0.5$ inches; and $w = 0.0833$ inches. When we test this dial, with exposure time, reading distance, etc., held constant, let us say that out of 100 readings there were five errors. From these data,

$$\frac{ls}{dw} = \frac{(0.5)\ (0.5)}{(12)\ (0.0833)} = 0.25$$

$$\frac{R_e}{R_t} = \frac{5}{100} = 0.05$$

This gives us one point of the desired curve. To obtain a second point, let us say that s is changed to 0.25 inches keeping the other values the same and that there were 15 errors in 100 readings. This would give an ls/dw of 0.125 and an R_e/R_t of 0.15. These and additional values can then be plotted to illustrate the function. Let us say that the plot showed a minimum value at $ls/dw = 0.3$, $R_e/R_t = 0.04$. This implies that the most efficient dials would have dimensions such that $ls/dw = 0.3$. If we wish to find the most efficient diameter for a dial of $s = 2$ inches, $l = 3$ inches, and $w = 1$ inch, we can substitute in the equation to give a d of 20 inches as the answer. Or if we have a dial of, say, $ls/dw = .15$ we can enter the curve to predict the proportion of errors for that dial when tested under the controlled conditions of the original investigation.

Perhaps our hypothetical investigations have also shown that the stroke-width of the markers has little influence on reading errors and that reading distance is a more important variable. Marker stroke-width can then become a controlled variable and reading distance a fundamental variable. The fundamental variables would be l, s, t and b (reading distance) and the ratios would be R_e/R_t, ls/db, etc.

It should be noted that the selection of a parameter is of considerable importance. There is an indefinite number of parameters containing the variables l, s, d and w and all are dimensionless: ls/dw; l/w — s/d; 210s/d + w/l; $(ls)^2/(dw)^2$; etc. The more complicated parameters must be considered when the simpler ls/dw does not yield a curve of suitable form (in the example it must go through a single minimum), or when the range of values of the ratio is so large as to be cumbersome. Frequently this selection is simplified by considering limitations of the variables and the parameters. For example, l is always less than d; 4ls/dt merely raises all values of the abscissa by a constant and does not change the shape of the curve; $(w/d)^5 + l/s$ is not suitable since the first part of the parameter is so small it loses significance. No matter which parameter is selected, the resulting curve will yield the same R_e/R_t for a dial of given geometry.

So far only four variables have been considered. Any number could have been selected as long as they could be reduced to one dimensionless parameter. For example, if l, s, d, b and w were considered as fundamental, the resulting parameter might be ls^2/dbw or lsd/bw^2. The procedure for obtaining data would remain unchanged with the exception of the increased freedom of including more variables. In this case, it would be possible to use only one gauge and change only the reading distance. The parameter would then change as well as the ratio of errors to readings so that it would be possible to plot the curve.

Dimensional analysis may thus be applied in any situation where there are independent dimensions such as in the selection of letter-size and spacing on highway signs, problems of room illumination, etc. Its efficient use of experimental time, simplicity of data analysis, and broad applicability of results recommend it as a useful technique.

REFERENCES

1. BUCKINGHAM, E. Model experiments and the form of empirical equations. *Trans. ASME,* 1915, **37**, 263-296.

2. LANGHAAR, H. L. Dimensional Analysis and Theory of Models. New York: Wiley, 1951.

506 Beech Avenue
Patton, Pennsylvania

Psychology Department
Lehigh University
Bethlehem, Pennsylvania

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 227-238.

AN EXPERIMENTAL QUANTIFICATION OF KINESTHETIC SENSATION: EXTENT OF ARM MOVEMENT*[1]

Institute for Psychological Research

PAUL G. RONCO

A. INTRODUCTION

The functional relation between subjective and stimulus magnitude has been shown to be a power function on an impressive variety of sensory continua (11). This has resulted in the development of a number of scales of sensory magnitude such as the "sone" scale of loudness, the "brill" scale for brightness and the "gust" scale for taste, to name but a few. However, scales of sensory magnitude such as found for other sense modalities have been especially lacking for kinesthesis.

It appears that the major portion of research on the modality has not been performed with the intent of investigating kinesthesis per se. Rather, the intent has very often been to solve some particular problem concerning man's ability to perform various types of movements under a number of environmental conditions. Of course, regardless of the intent, this wealth of data on motor abilities has increased our knowledge of kinesthesis.

In particular, studies of positioning movements have been fruitful in generating information about this sense modality (1, 13, 14). There are a number of reasons for this. A positioning movement is usually relatively simple in comparison to most other types of movements such as tracking or continuous reactions and this simplicity facilitates the measurement of various aspects of the movements. Also, their relative simplicity enhances the possibility of determination, control, and manipulation of the variables which affect such types of movements. When made without the aid of vision, positioning movements are primarily under the control of kinesthesis, and they therefore can be used as a rather direct measure of kinesthetic sensitivity.

The accuracy, for example, that an individual can attain in making a particular positioning movement is certainly a function of kinesthetic sensitivity. Furthermore, the various types of discriminations which can be made

---

* Received in the Editorial Office on October 8, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is part of a dissertation submitted to the Department of Psychology, Boston University, in partial fulfillment of the requirement for the Ph.D. degree.

in performing a positioning movement can give an indication of the types of dimensions along which kinesthesis can be scaled.

Resulting from the studies of positioning movement have been scales demonstrating the relationship primarily between such stimulus configurations as extent or direction of movement and usually such response measures as the constant or average error of movement. However, there have been developed three scales of subjective magnitude for continua that are at least in part mediated by kinesthesis; the "veg" scale for the subjective magnitude of heaviness, the "pac" scale for thickness or finger span, and a scale for the apparent magnitude of force. All of these scales are power functions (9, 11, 12).

The present study was directed toward the quantification of one aspect of kinesthetic sensation, namely the sensory attribute of extent of movement. A scale of subjective magnitude (the "kin" scale) was developed for this attribute through use of blind positioning movements using the psychophysical technique of magnitude estimation. In an attempt to substantiate the intended function and to determine which psychophysical method was most feasible for the quantification of this behavior the fractionation techniques of halving and doubling were also employed. In addition, a category scale was developed using the method of category production to ascertain its behavior when plotted against the scale of subjective magnitude, in an attempt to permit a preliminary classification of the resulting scale of sensory magnitude as either prothetic or metathetic according to the criteria presented by Stevens (10).

B. METHOD

1. *Apparatus*

The apparatus consisted of a 36-inch long track in the shape of a small I beam, attached to a plywood board, 26 inches wide, 36 inches long. A slider was attached to the track so that it fitted snugly yet could be moved up and down the track smoothly and easily. The top of the slider consisted of a plastic mold designed to accommodate the middle three fingers of an individual. Alongside the length of the track, at one-inch intervals, was a series of holes into which could be fitted a small peg. The insertion of the peg prevented movement of the slider beyond that point. Also running the length of the track was a rule, thus enabling the accurate measurement of the distance travelled by the slider. The plywood board was attached to a 30-inch high table which was bolted to the floor.

Also bolted to the floor facing the narrow end of the table was a chair, the seat of which was 18 1/2 inches from the floor. The apparatus board was so attached to the table that it could be slid toward or away from the chair without moving the table. When the board was moved in toward the chair, the end of the track was 10 inches from the seat reference point. This "end" point, which was quite close to the seated *S*'s body was the position from which all movements were started. There were straps attached to the seat so that the *S* could be strapped in, thus preventing the *S* from leaning forward while making a particular movement.

2. *Subjects*

The *S*s were twelve male college students between the ages of 20-30 years. Since it was considered a possibility that performance of the task might be somewhat affected by the anthropometric dimensions of the individual, the *S*s were equated for such things as height, weight, and anterior arm reach.

3. *Procedure*

a. General. The attempt was to construct a scale of subjective magnitude for extent of arm movement employing three different psychophysical techniques, namely magnitude estimation, ratio production (halving and doubling a standard stimulus), and category production. The procedure for each of these methods will be explained separately. There were, however, some procedural features which were common to all and which will be discussed first.

The basic task of the individual was in all cases first to make a blind positioning arm movement and then to make an instructed response on the basis of this movement, the particular response depending upon what psychophysical method was being employed at the time.

The *S* was strapped into the chair and blindfolded. The apparatus was always covered until the *S* was blindfolded so that at no time did the *S* see the apparatus. This was to minimize the possibility of an anchoring effect being developed from the *S*'s seeing the end points of the scale. The apparatus board was moved to the "in" position. The experimenter placed the *S*'s hand on the slide to get him in position and then read the instructions to him. In all of the techniques, magnitude estimation, category production, and ratio production the standard movement, i.e., the "extent" on which the *S* was to base his comparison judgment, was presented in the same manner.

The starting position for all movements was the same. The *E* placed the peg at the desired place along the track and instructed the *S* to move the slider as far along the track as possible, i.e., out to the peg. The *S* always moved out to the peg, back to the starting position, out to the peg, and then back again. In other words, he always made the standard or reference movement twice. By moving out to the stop twice the *S* obtained an idea of its position on the first excursion, so that on the second movement to the stop he did not have to move so cautiously as he frequently did the first time. Thus it was felt that rate of movement was controlled to some extent, at least wide differences in rate were controlled. This also prevented the *S* from running into the stop with great force, a factor which has been shown to have an effect on subsequent judgments concerning the movement (2). After the *S* came back to the starting position for the second time he carried out the instructed response depending on what part of the experiment he was performing. Each of the twelve *S*s went through all procedures, the order of presentation begins different for each *S*.

b. Magnitude estimation. In this section of the experiment using the technique of magnitude estimation the *S* made numerical estimates of the various movements, i.e., he assigned numbers to a finite series of stimuli under instruction to make the numbers proportional to the apparent magnitude of sensation. The *S* made a standard movement, seven inches in length and was told to call it 10. He then made a series of movements, their lengths being 1, 2, 3, 7, 11, 15 and 21 inches and assigned numbers to these movements. The *S* was given 21 trials (each movement presented three times in random order) in succession with no rest interval between trials. After the twenty-first trial a three-minute rest period was given. The standard was presented again and 21 more trials were given. These 42 trials constituted one session. The next day the procedure was repeated. In all the *S* was given 84 trials of magnitude estimation comprised of 12 judgments for each of the seven stimulus extents.

c. Ration production. The task of the *S* with this procedure was either to halve or double a standard movement of various extents.

For one procedure the *S* was required to halve movements of 1, 3, 7, 11, 15 and 21 inches in length. Each of these movements was presented to the *S* seven times in random order. The 42 trials comprised the sessions.

For the other procedure, the *S* was required to double, i.e., make movements twice the length of standard movements of 1, 2, 3, 7, and 11 inches in length. He made seven judgments of each of the five lengths for a total of 35 trials. The order of stimulus presentation was random.

d. Category production. The basic task of the *S* with this procedure was to make movements which he felt belonged to one of seven equally spaced categories, after he had experienced the extreme categories. The *S* was given a two-inch movement and a 19-inch movement and was told that these movements belonged to categories "1" and "7" respectively. They were then asked to make movements which belonged to one of seven equally spaced categories. For example, the *E* would say "seven" and the *S* would make a movement which he felt belonged to the seventh category. The categories were presented in random order. The *S* was given each of the seven categories three times for a total of 21 successive trials. He was then given a three-minute rest after which the two standard movements were presented again and then 21 more trials. This constituted one session. The same procedure was given again the next day. The *S* was asked to produce each of the seven categories 12 times for a total of 84 trials.

C. RESULTS AND DISCUSSION

The results of the magnitude estimations are presented in Figure 1 with the median magnitude estimation plotted against the reference movement in inches. Each point on the curve represents 144 estimates or 12 estimates each by the 12 *S*s. The vertical lines represent the interquartile ranges of the estimates. The plot of the data was obtained by the method of least squares.

The median estimates follow a straight line on a log-log plot indicating that the apparent or subjective extent of arm movement is related to the physical distance by a power function, i.e., it conforms to the equation $K = cI^n$ when $K =$ sensory magnitude, I is the physical distance in inches, and c is a constant. The slope of the line which gives the exponent of the power function is 1.05. Thus, subjective magnitude of extent of arm movement is a positively accelerated function of stimulus magnitude (physical extent of movement) and this equation holds over the entire range of movement lengths examined.

The median results of each individual *S* can be fitted equally as well by a straight line in a log-log plot indicating that the power function relation holds for each *S* and that the function in Figure 1 is not an artifact that depends upon the averaging of the individual data. From an examination of the interquartile ranges it can be seen that the absolute variability increases with increasing distance but the relative variability decreases. This is typically the case when the task of the *S* is to duplicate or fractionate an extent of movement.

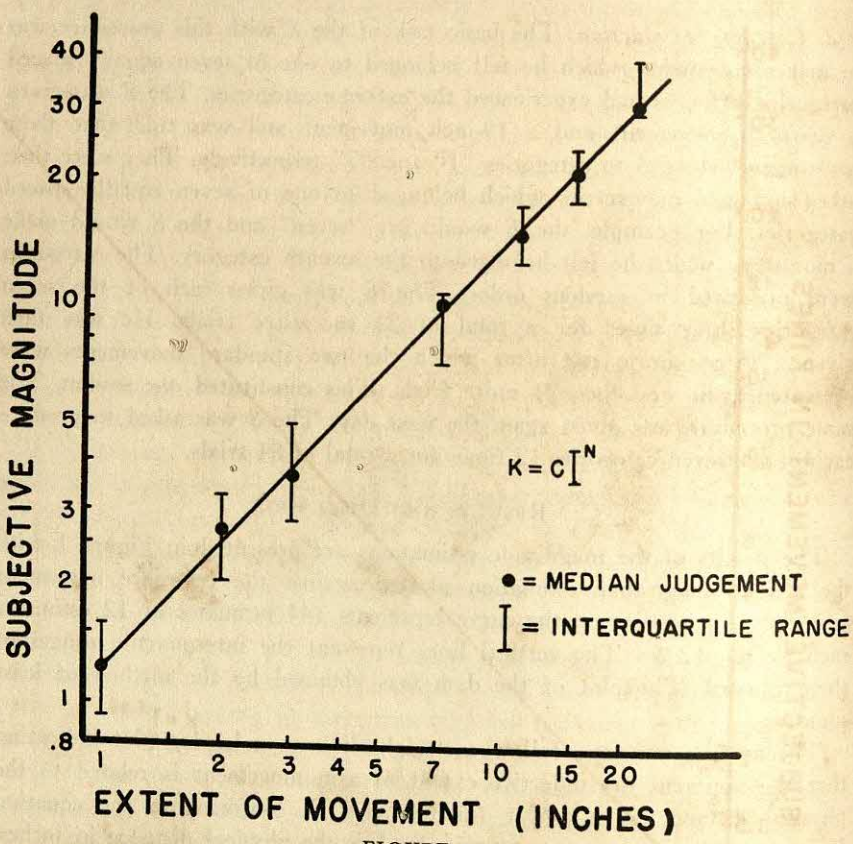


FIGURE 1
MEDIAN MAGNITUDE ESTIMATION OF EXTENT OF MOVEMENT IN THE HORIZONTAL MEDIAL PLANE

The curve in Figure 1 was used to construct a scale of subjective movement. The unit of this scale is called the Kine from the term kinesthesis and represents the subjective magnitude of a movement one inch in length. The relation of Kines to inches of movement is presented in Figure 2. The relation can be expressed by $K = I^{1.05}$ where K is Kines and I represents inches of movement. It can be seen from the Kine scale that equal stimulus ratios produce equal subjective ratios. A ratio of 1:1.94 on the stimulus scale produces a subjective ratio of 2:1 and this relationship remains constant over the entire extent of the scale.

Figure 3 presents the median judgments of the ratio production procedures. The solid upper line represents a 2 to 1 relation between the

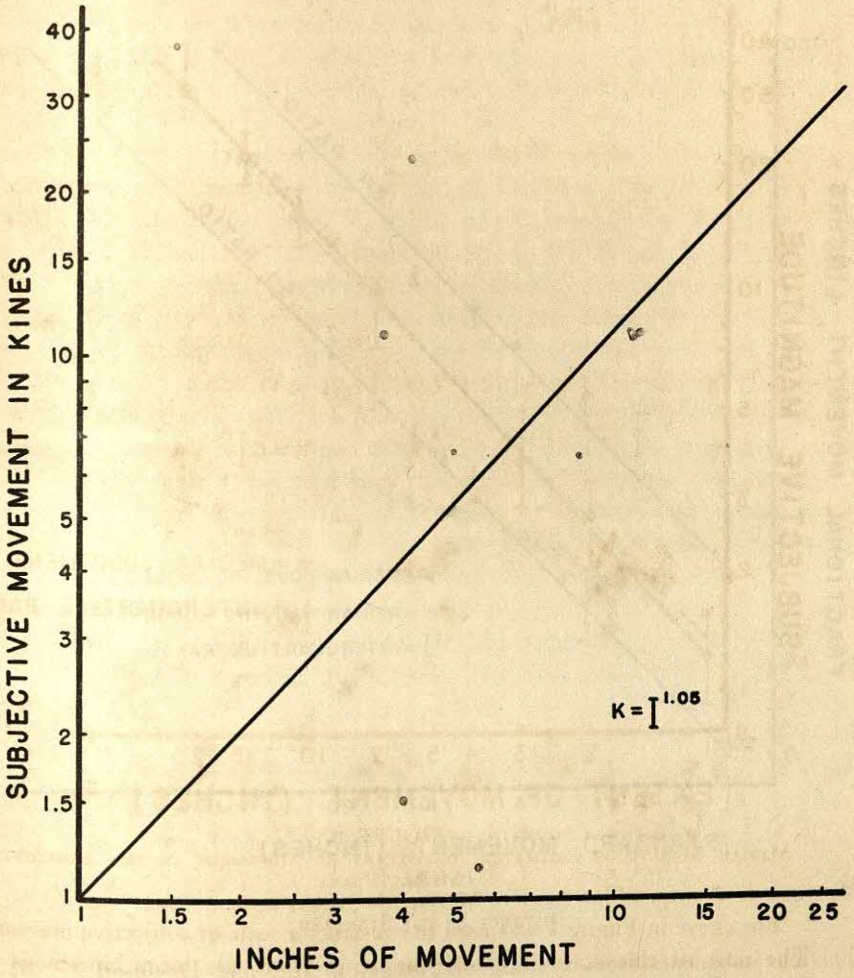


FIGURE 2
KINE SCALE OF SUBJECTIVE EXTENT OF MOVEMENT

variable and the standard movement in inches. This line can be used to evaluate the doubling settings made by the *S*. The lower line represents a .5 to 1 relation between the variable stimulus and the standard and can be used as a base against which to evaluate the halving judgments. The slopes of these lines are 1.

As can be seen, the halving judgments can be fitted quite well to a straight line on a log-log plot. The X's along the regression line represent

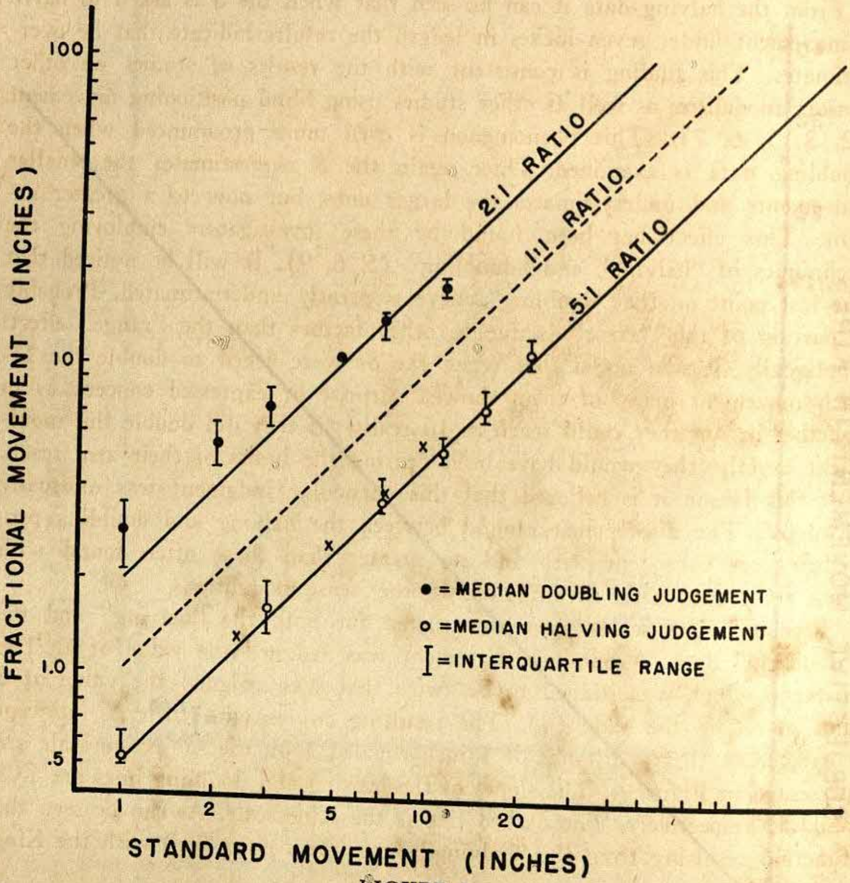


FIGURE 3
RELATION OF STANDARD MOVEMENT TO MOVEMENTS JUDGED TO BE 2 TIMES AND 1/2 THE DISTANCE OF THE STANDARD

halving judgments from a study of Spragg, Devoe, and Davidson (7, 8). Despite the fact that their task was somewhat different than the one in the present experiment, there is a high degree of consistency between the two sets of data. The doubling data, however, does not seem to follow a straight line. The point at 5 inches was derived from an earlier experiment and represents the median of 24 doubling estimates made by six *S*s. It is put in here simply to fill in the function. Under the conditions of this experiment it can be seen that *S*s perform more accurately in terms of adjusting to the prescribed ratio by the operation of halving than by doubling. The median plots for halving are very close to the "ideal" line.

From the halving data it can be seen that when the *S* is asked to halve a movement under seven inches in length the results indicate that he overestimates. This finding is consistent with the results of studies on other sensory modalities as well as other studies using blind positioning movement (2, 3, 5, 6, 7). This phenomenon is even more pronounced when the doubling data is examined. Once again the *S* overestimates the smaller movements and underestimates the larger ones, but now to a greater extent. This effect has been found by these investigators employing the techniques of "halving" and "doubling" (5, 6, 9). It will be noticed that the last point on the "doubling" curve is greatly underestimated. Probably a portion of this "error" is due to other factors than the "range" effect. Specifically, it was noted that when the *S*s were asked to double the 11-inch movement many of them showed surprise or expressed concern as to whether or not they could reach it. In reality, if they did double this movement exactly, they would have been nearing the limits of their arm reach. For this reason it is believed that this particular judgment was unusually distorted. The discrepancies found between the halving and doubling procedures are consistent with and no greater than those often found with these techniques in the investigation of other sense modalities.

Psychophysical scales were constructed for both the "halving" and the "doubling" data. A one-inch movement was assigned the value of 1. The distance which was judged to be twice that was assigned the value of 2 and so on up the scale (4). The resulting curves with their Y intercept corrected so that a distance of 1 inch equaled 1 on the subjective scale are presented in Figure 4. The slopes of the halving and doubling lines are 1.08 and .78 respectively. The dashed line is the Kine scale. As can be seen the function resulting from the "halving" data agrees quite well with the Kine scale.

The function obtained from the doubling data, however, appears to be quite different in slope from that obtained by the other procedures. It is noteworthy that most *S*s did not seem as confident in making this judgment as they did in making the others. Most of them remarked about this procedure being more difficult. Whatever the reasons, the discrepancy between the curve derived from the doubling data and those derived from the other techniques seems large enough to warrant skepticism concerning the use of the present doubling data for the purpose of scaling subjective magnitude.

Figure 5 presents the results of the category production procedure plotted against the scale of subjective movement in Kines. Each point on the curve represents 144 judgments. As can be seen, the plot is linear, indicating that

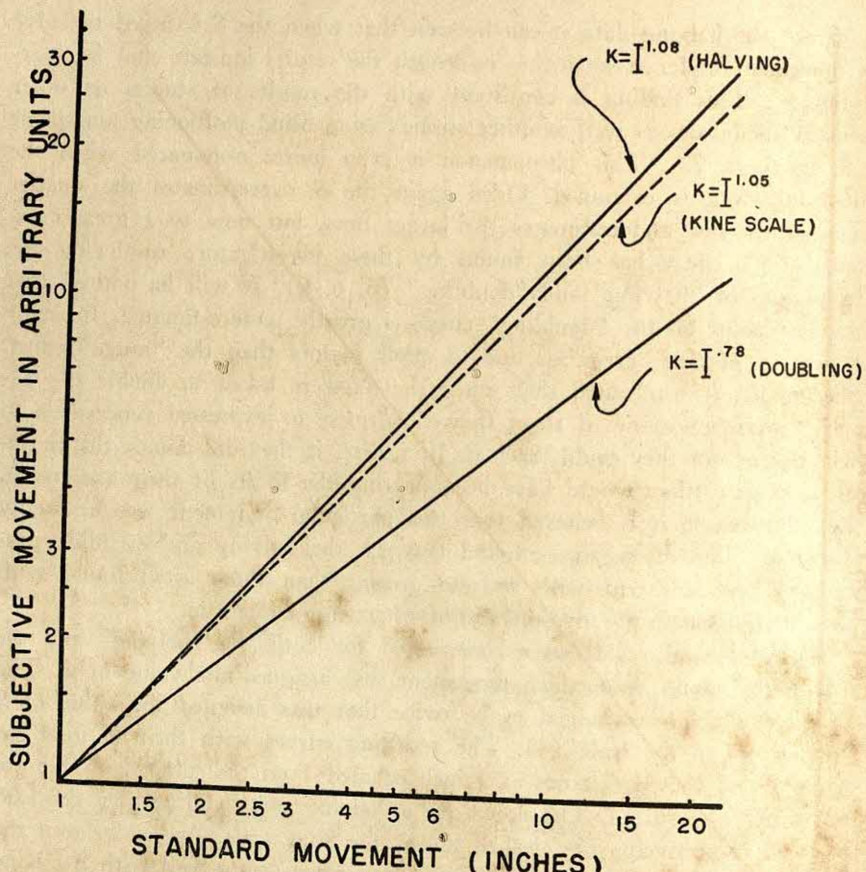


FIGURE 4
SUBJECTIVE SCALES OF MOVEMENTS BASED ON RATIO PROCEDURE: HALVING AND DOUBLING

discriminal sensitivity is comparatively uniform over the lengths of the scale. While not conclusive, this, along with the fact that the subjective extent of movement grows as a function of physical movement at an only slightly accelerated rate, suggests the extent of movement may possibly be classified as a metathetic continuum. However, it should be kept in mind that the criterion for metathetic is not always as clear as that for prothetic. The present uncertainty as to whether or not we are dealing with a metathetic continuum arises in part from a lack of knowledge concerning the size of the JND for extent of arm movement. If there were data showing that the subjective size of the JND remains fairly constant over the extent of the scale, then it could be said with somewhat more confidence that

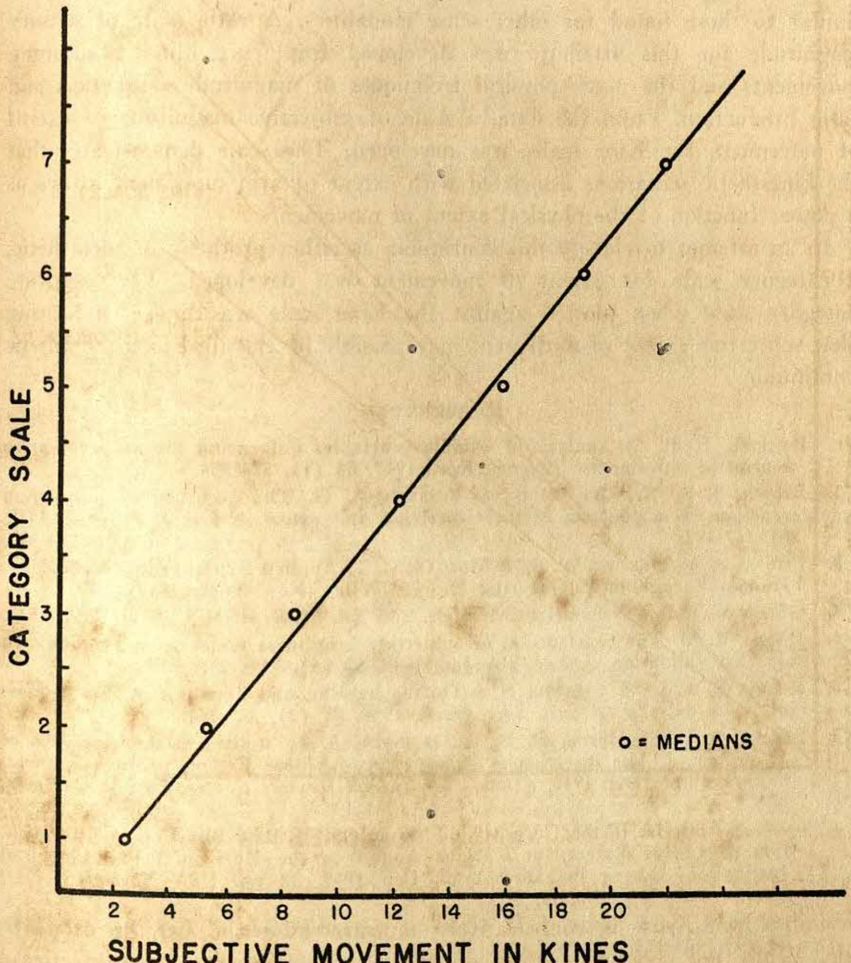


FIGURE 5
CATEGORY SCALE PLOTTED AGAINST THE SCALE OF SUBJECTIVE MOVEMENT IN KINES

extent of movement is a metathetic continuum. In order for the subjective size of the JND to remain constant over the range of this continuum, the absolute size of the JND would have to remain about the same. This can be seen from the Kine scale.

D. SUMMARY

An attempt was made to quantify an attribute of kinesthetic sensation, namely extent of movement, and to construct a scale of sensory magnitude similar to those found for other sense modalities. A ratio scale of sensory magnitude for this attribute was developed employing blind positioning movements and the psychophysical techniques of magnitude estimation and ratio production. From the data, a scale of subjective magnitude of extent of movement, the Kine scale, was developed. The scale demonstrates that the kinesthetic sensations associated with extent of arm movement grows as a power function of the physical extent of movement.

In an attempt to classify this continuum as either prothetic or metathetic, a category scale for extent of movement was developed. The resulting category scale when plotted against the Kine scale was linear, indicating that subjective extent of movement may possibly be classified as a metathetic continuum.

REFERENCES

1. BAHRICK, H. P. An analysis of stimulus variables influencing the propreoceptive control of movements. *Psychol. Rev.,* 1957, **64** (1), 324-328.
2. BROWN, J. S., KNAUFT, E. B., & ROSENBAUM, G. The accuracy of positioning reactions as a function of their direction and extent. *Amer. J. Psychol.*, 1948, **61**, 167-182.
3. CHAPANIS, A., GARNER, W. R., & MORGAN, C. T. Applied Experimental Psychology: Human Factors in Engineering Design. Wiley, New York, 1949.
4. GUILFORD, J. P. Psychometric Methods. 2nd Ed. McGraw-Hill, New York, 1954.
5. HANES, R. M. The construction of subjective brightness scales from fractionation data: a validation. *J. Exp. Psychol.*, 1949, **39** (5), 718-728.
6. POULTRAN, F. C., & STEVENS, S. S. On the halving and doubling of the loudness of white noise. *J. Acoust. Soc. Amer.*, 1955, **27** (2), 329-331.
7. SPRAGG, S. D. S., DEVOE, D. B., & DAVIDSON, A. L. Studies in the Accuracy of Movement. I. The Bisection of Duplication of Linear Extents in the Horizontal Sagittal Plane. Oct. 1949, 20 pp. *USN Special Devices Center*, Port Washington, N.Y.
8. ———. Studies in the Accuracy of Movement. II. The Bisection and Duplication of Linear Extents at Various Angles in the Horizontal Plane, and at Inclinations Above the Horizontal. Oct. 1949, 22 pp. *USN Special Devices Center*, Port Washington, N.Y.
9. STEVENS, J. C., & MACK, J. D. Scales of apparent force. *J. Exp. Psychol.*, 1959, **58** (5), 405-413.
10. STEVENS, S. S. On the psychophysical low. *Psychol. Rev.*, 1957, **64** (3), 153-181.
11. STEVENS, S. S., & GALANTER, E. G. Ratio scales and category scales for a dozen perceptual continua. *J. Exp. Psychol.*, 1957, **54** (6), 377-411.
12. STEVENS, S. S., & STONE, G. Finger span: ratio scale, category scale, and JND scale. *J. Exp. Psychol.*, 1959, **57** (2), 91-95.
13. WEISS, B. The role of propreoceptive feedback in positioning responses. *J. Exp. Psychol.*, 1954, **47** (3), 215-224.
14. WEISS, B., & SPRAGG, S. D. S. Building "Feel" into Controls: The Effects on Motor Performance of Different Kinds and Amounts of Feedback. Oct. 1953, 20 pp. *USN Special Devices Center*, Port Washington, N.Y.

Institute for Psychological Research
Tufts University
Medford 55, Massachusetts

$20.00 per annum
$15.00 per volume
$7.50 per half volume

QUARTERLY
Two volumes per year
Immediate publication

April, 1963
Volume 55, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

APRIL, 1963

(Manuscripts are printed in the order of final acceptance)

A new automatic circuit for the Selective Mathometer (MK II) 241
By Clyde E. Noble, James E. Fuchs, and Jack D. Thompson

Transitional probability in anagram solution in a group setting . . . 259
By Thomas Stachnik

Prior parochial vs. secular secondary education and utterance latencies to taboo words 263
By George S. Grosser and William J. Laczek

Development of patriotism in children—A second look 279
By Edwin D. Lawson

Temporal judgment as a function of nosological classification and experimental method 287
By Joel S. Warm, James R. Morris, and John K. Kew

(OVER)

Copyright, 1963, by The Journal Press
Provincetown, Massachusetts, U. S. A.
Entered as second-class matter January 26, 1937, at the post-office at
Provincetown, Mass., under the Act of March 3, 1879
Second-class postage paid at Provincetown, Mass.

The relationship between persistence and personality 299
By H. M. Brachman and C. G. Costello

Stuttering and self concept 307
By Robert W. Rieber

The disadvantaged group: A concept applicable to the handicapped . . 313
By Sidney Jordan

Lateral dominance characteristics and reading achievement in the first grade . 323
By Irving H. Balow

Family constellation and schizophrenia 329
By Arthur Burton and Joseph W. Bird

The effect of mediators in concept identification 337
By Daniel S. Lordahl

Pain sensitivity of amputation extremities 345
By Hyman Korin, Samuel A. Weiss, and Sidney Fishman

The effects of small and large group instruction on learning of subject matter, attitudes, and interests 357
By John F. Feldhusen

Measurement of palmar sweating: Effect of repeated measurement from the same area 363
By H. Carl Haywood and Donald J. Shoemaker

Experimental evidence for the involvement of a neurophysiological mechanism in the discrimination of duration 371
By Thomas M. Nelson, S. Howard Bartley, and James F. Jordan

The brightness of part-spectrum targets stimulating the eye parafoveally at intermittency rates below fusion 387
By Thomas M. Nelson, S. Howard Bartley, and Zachary Ford

Daily preferences and institutionalization 397
By Harry F. Desroches and Charles L. Walter

Combining external and internal symbolization in the LSD episode . . . 401
By Donald K. Gucker

The retention of incidentally learned connotative responses 409
By Halide S. Yavuz

Language processing ability of process and reactive schizophrenics . . . 419
By David Pearl

The effect of voluntarily directed attention on successive cortical activation responses 427
By Thomas Mulholland and Sylvia Runnals

A study of self-disclosure in the classroom 437
By Philip Himelstein and Wilson W. Kimbrough, Jr.

Fakability of the Gordon Personal Inventory: Replication and extension . . 441
By John R. Braun

Personality factor stability for three ability levels 445
By Gerald H. Shure and Miles S. Rogers

A method for detecting subgroups in a population and specifying their membership 457
By J. A. Gengerelli

Anagram solution times: A function of word length and letter position variables 469
By M. S. Mayzner and M. E. Tresselt

Classification transfer with disjunctive concepts as a function of verbal training and set 477
By Hans G. Furth

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 241-257.

A NEW AUTOMATIC CIRCUIT FOR THE SELECTIVE MATHOMETER (MK II)*

Department of Psychology, Montana State University

CLYDE E. NOBLE, JAMES E. FUCHS, AND JACK D. THOMPSON[1]

A. INTRODUCTION

The Selective Mathometer (Gr. *manthano* = to learn + *metron* = measure) is an electro-mechanical serial pushbutton apparatus for experimental research in human multiple-choice learning. The prototype (MK I) was originally designed by the senior author in 1952 at the U.S. Air Force Human Resources Research Center. It was described in this Journal in 1955 (3). Mark I was a semiautomatic instrument of simple construction yet versatile operation which combined the essential features of the maze and the memory drum in a novel manner. The Mathometer was developed to meet the need for a flexible laboratory device of analytic design and high reliability which could serve as a standard reference task for systematic investigations of a wide range of learned behaviors of the probabilistic, sequential type. This class of *selective learning* situations (e.g., trial-and-error, serial, discrimination learning problems) does not ordinarily require gradations in latency, amplitude, force, or precision of action as do such traditional perceptual-motor skills as dexterity, coordination, and tracking. We are, therefore, less interested in the fine grain of so-called "response learning" than in the molar associative processes whereby individual S-R links and serial action chains are formed. As the Bilodeaus expressed it in a recent review of motor skills, "Devices such as Noble's Mathometer deliberately reduce the R component close to zero" (1, p. 246). A general description of our research program was published in 1957 (2), a reference which still provides a useful introduction to the aims and methods of this laboratory.

---

* Received in the Editorial Office on October 8, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This research was partially supported by a grant to the first author from the National Science Foundation. The second author has served as principal NSF assistant in the Montana Laboratory since 1958. The third author, an electrical engineer, is now with the Boeing Aircraft Corporation in Seattle, Wash. Valuable assistance in various phases of the developmental work was rendered by Blaine L. Baker, Ridgely W. Chambers, Frederick A. Hatfield, Roland J. Keefer, Frederick W. Krieger, Richard W. Peterson, Shirley Quanbeck, Merle Ridgley, and Robert H. Scherpenseel.

The major difference between the present apparatus and its semiautomatic predecessor is that the degree of automation has been greatly increased during the past decade. All temporal events, reward functions, and pacing variations are now automatically controlled through the use of selector switches, decade-type electronic interval timers, and a master stepping switch. Permanent graphic records of the sequence, duration, and latency of all responses are collected on a multichannel operations recorder, response frequencies being verified at the end of each trial by magnetic digital counters. In addition, measurements of stimulus exposure duration, inter-stimulus (dark) period, reward delay, and reward duration are periodically monitored by synchronous clocks to assure calibration constancy throughout an experimental session. The final touch in environmental control has been brought about by air conditioning and sound deadening of the laboratory. As a result, we now have the capability of both audio and visual isolation of *S* while maintaining continuous one-way observation, thereby increasing the precision and the potential significance of experimental variables which may be introduced.

B. General Description

During the summer of 1961 the Department of Psychology at Montana State University occupied new quarters in a permanent three-story building of brick and stone construction on the central campus. Renovated to accommodate the Department's plans to expand its graduate training facilities, the Psychology Building provides adequate space for a variety of research, teaching, and other professional psychological activities. The Selective Learning Research Laboratory is located in the east wing of the first floor, an area measuring approximately 24 feet × 50 feet, 154 square feet of which are divided into a pair of sound-deadened, air-conditioned chambers designed and equipped for research in mathometry. The remaining 1046 square feet in this wing are devoted to data collection and analysis of studies in verbal behavior and psychomotor skill. A floor plan of the laboratory is shown in Figure 1. The major equipment components are: *S*'s response panel and projection viewscreen, identical to those reported earlier (3); automatic slide projector with mirrors for through-wall projection; operations recorder; *E*'s control unit; Timer-Relay cabinet; and the intercommunication system. For sound deadening, the walls, ceiling, and doors are covered with 3/4 inch acoustic tile, floor coverings are in cork, electrical outlets are filled with spun glass, wall and floor joints are caulked and covered with moulding, and the doors are stripped with foam rubber gaskets. During experimentation the masking of extraneous noise is accomplished through the use of a 1/2 ton

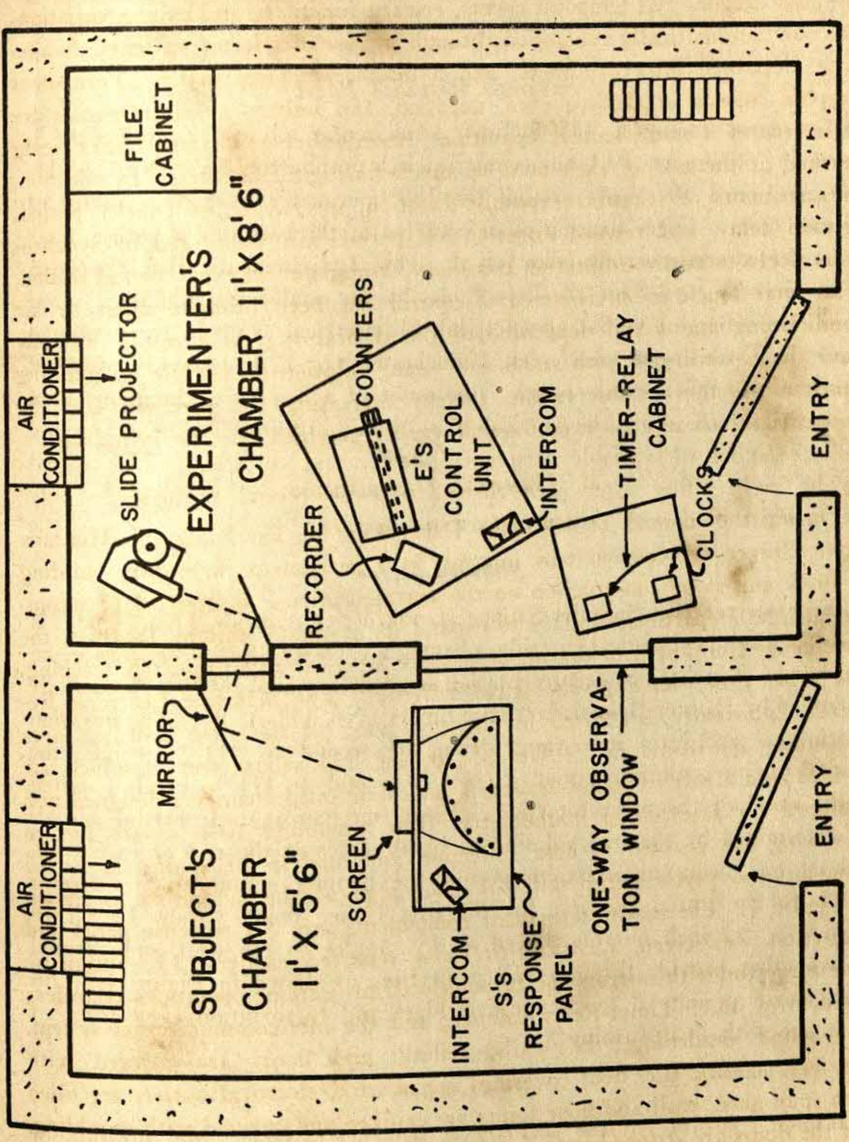


FIGURE 1

Floor Plan of the Selective Learning Research Laboratory air conditioner in *S*'s chamber and a one ton unit in *E*'s chamber. A one-way mirror and lighting differential between the rooms allows *E* to observe *S* at all times, and *S* is instructed and monitored via a two-way intercommunication system.

1. *Subject's Response Panel*

S is seated facing a 45° inclined semicircular plywood panel (cf. 3, Figure 1). There are 19 Unimax microswitch pushbutton keys (No. MXJ1-728) positioned 10° apart around the 180° perimeter, each at a radius of one foot from a finger fixation point centered at the base of the panel. Keys are numbered consecutively from left to right. Individual detachable wooden blocks serve as covers for the keys *E* decides to render unavailable. Visual stimuli consisting of verbal, geometrical, or electrical symbols are presented serially on a translucent viewscreen, 10 inches high by 12 inches wide, mounted directly above the response panel. The onset of a 3-w green lamp installed below the screen along the vertical axis of Key 10 provides a reinforcing feedback signal of variable frequency, delay, and duration. This signal may be given either after each correct response (serial reward) or only after a series of correct responses (teminal reward).

2. *Slide Projector*

The projector is a Spindler-Sauppe standard Selectroslide (No. 300). It is equipped with cylindrical cartridges having a capacity of 48 slides measuring 2 inches × 2 inches. Slide presentation is predetermined by *E* and remotely controlled by Hunter-Brown electronic timers (No. 111-A, B). The duration of stimulus exposures may range from .05 second to 111.0 seconds, the interval for inter-stimulus periods from 1.70 seconds to 111.0 seconds (limited by the speed of the projector slide-changing mechanism). Intertrial periods are determined by the interpolation of blank slides at the end of each trial, at least one being necessary to permit the homing action of the selector switch to the initial position for *S*'s first choice point (Slide 1). Two 12 inch × 12 inch mirrors placed at 45° angles to the wall and parallel to each other provide through-wall projection, as shown in Figure 1. This arrangement permits *E* to change cartridges and to maintain good focus of the image without disturbing *S*.

3. *Recorders*

Graphic recording of the occurrence, sequence, latency, and duration of all responses as well as stimulus presentation and reward signals is provided by a 20-pen Esterline-Angus Operations Recorder (Model AW). Available chart speeds are 3/4, 1-1/2, 3, 6, and 12 inches per hour or per minute, and 3/4, 1-1/2, and 3 inches per second. Our typical rate is six inches per minute. Two six-digit Veeder-Root electromagnetic Digital Counters (No. 1205) also cumulate the number of correct (R +) and total (R) responses which are collated with the graphic recordings in final data analyses.

4. *Experimenter's Control Unit*

A view of *E*'s chamber is shown in Figure 2. The control unit is housed in a 12 inch × 18 inch × 12 inch metal cabinet. On the top face of the

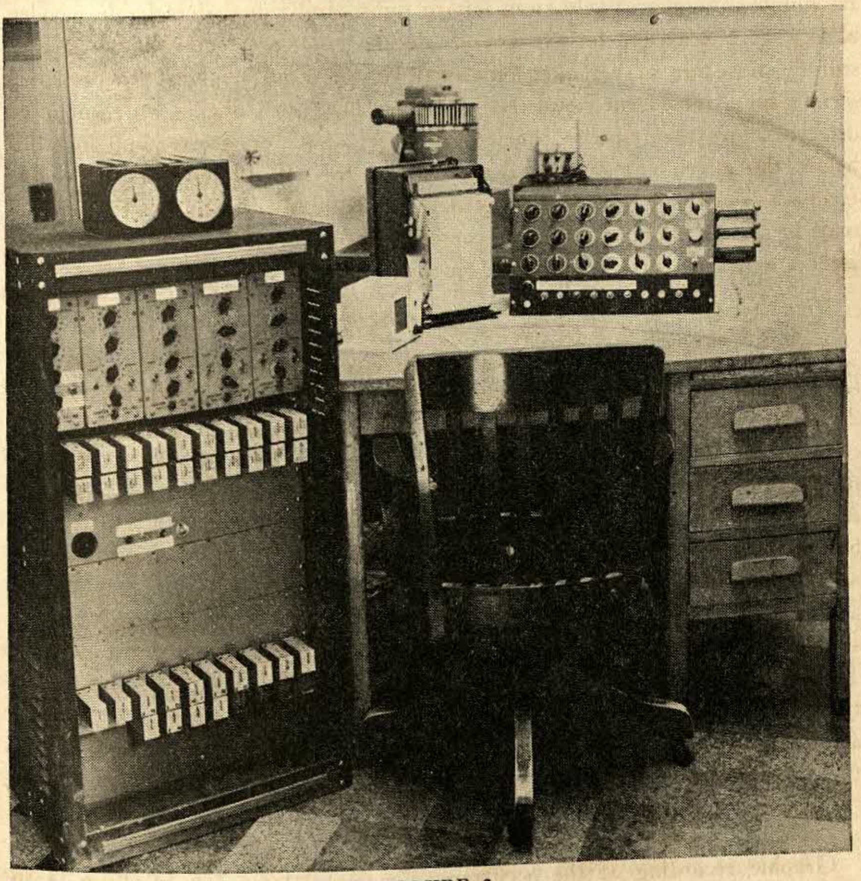


FIGURE 2

VIEW OF *E*'S CHAMBER

From left to right the major items of equipment are: one-way screen, timer-relay cabinet, intercom, operations recorder, projector, *E*'s control unit.

control panel are 18 rotary switches which correspond to the 18 reaction keys on *S*'s response panel. The position of each rotary switch establishes a circuit relationship between a pushbutton on the response panel and a stimulus slide. For example, if *E* desires that Key X ("X" signifying any integer from 1 to 18) on *S*'s panel should be the correct response for the first symbol on the screen, then Switch No. 1 must be set to position No. X. Similar settings must be made for all other stimulus-response (slide-key) connections. A manually-operated pushbutton reward switch and a reward indicator lamp are also located beside the rotary switches for *E*'s use. At the bottom of the panel arranged from left to right are the following seven switches:

(*a*) Power Switch, which applies 120 v. AC throughout the entire system.

(*b*) Projector Switch, which provides power to the projector and the stepping relay control.

(*c*) Pacing Switch, which permits either automatic *E*-paced control of the slides (up position) or semiautomatic *S*-paced control (down position), depending jointly on *S*'s response and the position of switch S31 in the Timer-Relay cabinet (to be described later).

(*d*) Self-Pacing Criterion Switch, which is non-functional in the present system.

(*e*) Time Criterion Switch, which either requires *S* to respond while a stimulus is on the screen (up position) or allows him to respond at any time prior to the onset of the next slide (down position).

(*f*) Serial-Terminal Reward Switch, which permits reward either after each correct response (up position) or after the termination of a series of correct responses (down position). Reward may then be given either manually by *E* or the switch may be closed prior to the last response in the sequence; in this latter case, the final response in the sequence determines whether the series is rewarded or not.

(*g*) Manual Advance Switch, which allows *E* to advance the stimulus sequence at any time during *S*-pacing.

Also located on *E*'s control panel and visible in Figure 2 are a power indicator lamp and the two Veeder-Root counters which cumulate frequency of correct (R+) and total number of responses (R).

5. *Timer-Relay Cabinet*

The major portion of the new automatic circuitry for the Selective Mathometer is contained in the Timer-Relay cabinet, also pictured in Figure 2. The five Hunter-Brown electronic timers control stimulus duration, inter-stimulus period, reward delay, reward duration, and type of pacing. Calibration of the four time variables in units of .01 second is achieved independently by two modified Standard Electric time clocks (Model S-1). Typical values for these constants in most of our experiments have been two seconds, 1.70 seconds, 0 second, and .75 second respectively. The functional interdependence of the individual units is illustrated in Figure 3. Cable CE connects *E*'s control unit with the Timer-Relay cabinet, Cable CR connects the Timer-Relay cabinet with the operations recorder, projector control is maintained through cable CP, and *S*'s response panel is connected through cable CS.

6. *Intercommunication System*

A Knight intercommunication system (No. 83Y295) has been installed so that the master unit is in *E*'s chamber (cf. Figure 2) and the remote station is directly to the left of *S*'s response panel. This type of intercom is desirable because only the master station needs to be connected to the AC power outlet. If the remote station is secured in the "talk" position *E* may monitor *S* at all times, but *S* cannot receive the master station unless *E* places the master unit in the "talk" position.

C. CIRCUIT FUNCTIONS

The major new improvement in the automatic circuitry is the addition of a telephone type stepper-relay system (labeled R26A to R26E) which is schematized in Figure 4. In the former semiautomatic apparatus (MK I) the reward sequences had to be advanced manually by *E*, a task requiring considerable skill and sustained vigilance on *E*'s part. This function is now performed automatically by Hunter-Brown timers, the stepping switch (Western Electric No. 162757), and the relay matrix, producing greater dependability and freedom from experimenter-error. Mathometer reliability coefficients, based on odd-even trial correlations from several hundred *S*s (corrected for attenuation by the Spearman-Brown formula), have been found to range from .94 to .99 depending upon task factors (number of responses) and type of criterion score (R+ vs. R—) employed (4). Counter records agree closely with graphic records, but the latter are needed for analyzing initial R+ scores, for sequences, and for measuring temporal aspects of performance.

During a standard *E*-paced experiment the reward sequence is determined by the positions of the rotary switches of *E*'s control unit diagrammed in Figure 5. Each contact of Section E of the stepper relay is connected to the center arm of one of the 18 rotary switches. Thus each rotary switch

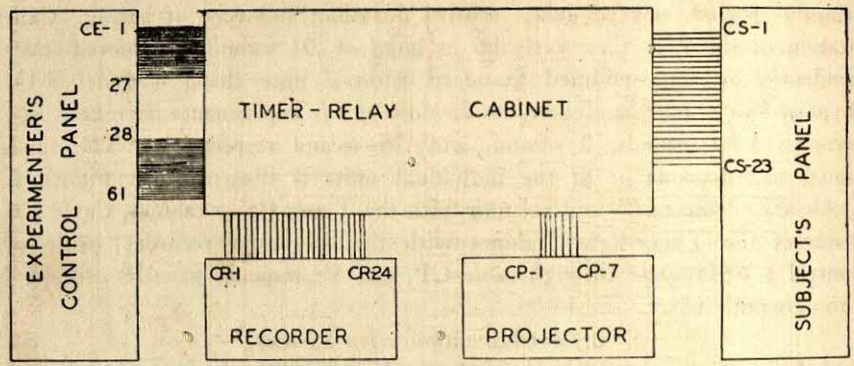


FIGURE 3
CABLE CONNECTIONS BETWEEN *E*'S CONTROL PANEL, *S*'S RESPONSE PANEL, RECORDER, AND PROJECTOR THROUGH THE TIMER-RELAY CABINET

represents one stepper position, and each rotary switch may be set to any of 18 positions, each position corresponding to one of the 18 functioning keys on *S*'s response panel as seen in Figure 6. Therefore, any combination of stimulus, serial position, and response key position can be achieved. To illustrate, if rotary switch No. 1 is set to position No. 7, Response Key 7 will be correct for the first slide. In a similar manner, each consecutive rotary switch is positioned to correspond to one of the slides in the stimulus sequence. It should be noted that a particular key may be designated as correct for more than one slide. Response Circuit No. 2 has been temporarily deleted, and Esterline-Angus pen No. 2 now records duration of reward. Pen No. 20 records stimulus onset, duration, and termination.

The rate of stimulus presentation is predetermined by the settings of Timer No. 3 (stimulus duration) and Timer No. 4 (inter-stimulus interval) during *E*-paced experiments. To start a sequence of trials *E* closes the projector switch on his control panel, thus activating a pair of microswitches in the internal circuitry of the Spindler-Sauppe projector and initiating the delay cycle of Timer No. 3. The impulse from Timer No. 3 energizes relay R20, the projector lamp, and the stimulus duration pen of the Esterline-Angus recorder. Timer No. 4 controls the duration of the dark period between stimuli, and begins operating the moment Timer No. 3 completes its timing cycle. At the beginning of the dark period, 120 v. AC power is applied through relay R24 to the slide change motor in the projector as diagrammed in Figure 7. This pulse moves the slide cartridge to the next stimulus position. The circuit to the change motor is then mechanically

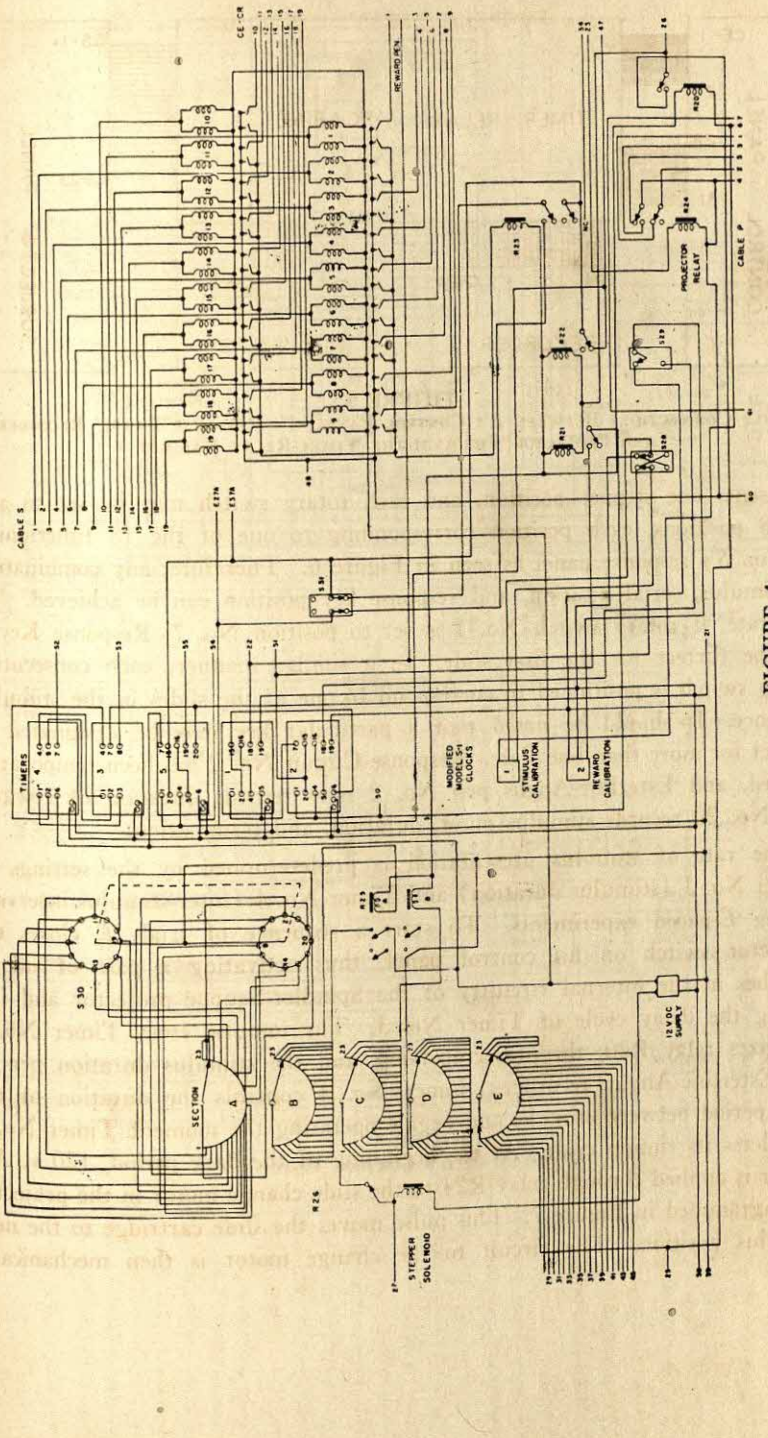


FIGURE 4

SCHEMATIC DIAGRAM OF THE STEPPER-RELAY REWARD-PROGRAMMING SYSTEM

For simplicity the internal schematics of the timers are omitted, together with the Jones plug circuits in *S*'s panel, *E*'s panel, and in the relay panel.

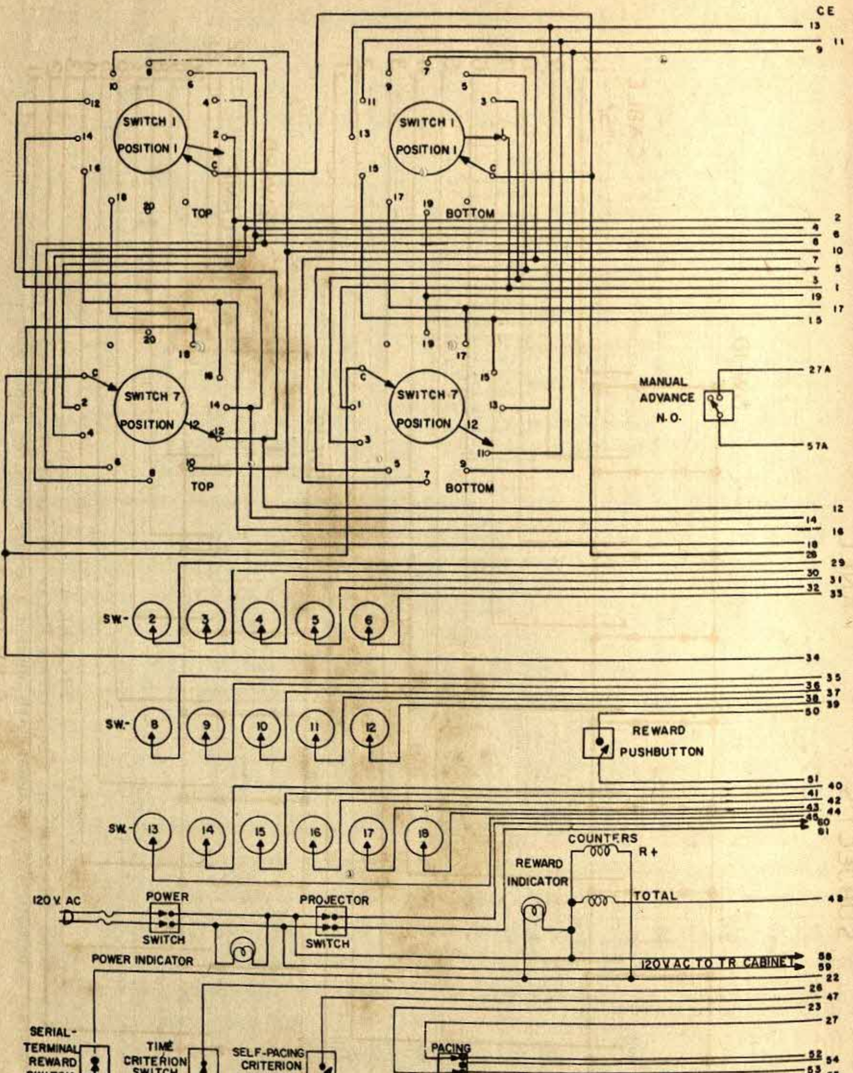


FIGURE 5
DIAGRAM OF *E*'S CONTROL UNIT

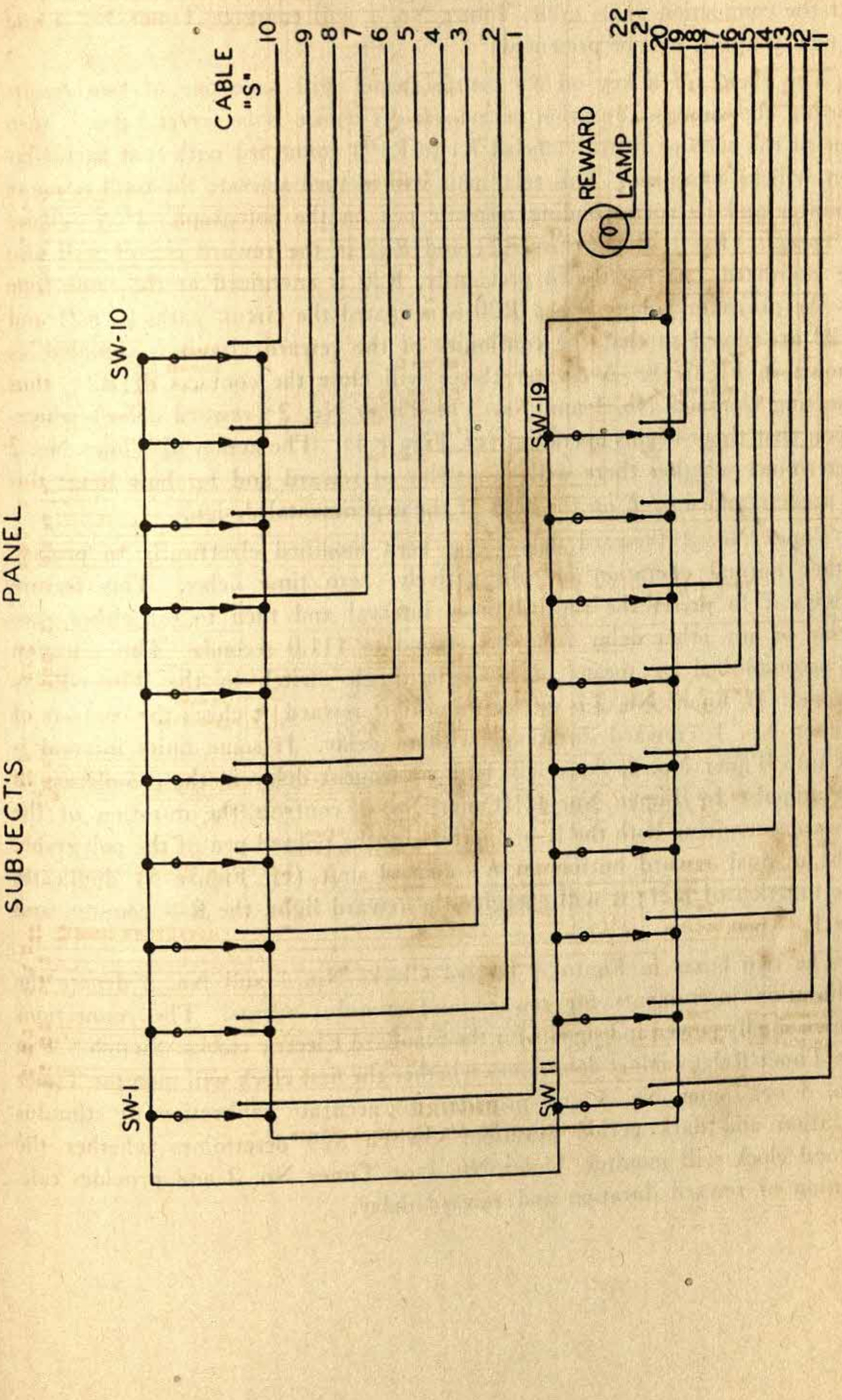


FIGURE 6
DIAGRAM OF *S*'s RESPONSE PANEL

opened by a microswitch which also closes the path to the projector lamp. At the completion of its cycle, Timer No. 4 will energize Timer No. 3 and a new stimulus will be presented.

The choice of a key on *S*'s control panel will have one of two results during the stimulus duration period. If *S*'s choice is *incorrect* (R—), then one of the pairs of relays (labeled R1 to R19) connected with that particular key will be energized. This relay pair will in turn activate the total response counter and its corresponding response pen on the polygraph. If *S*'s choice is *correct* (R+), then relays R21 and R22 in the reward circuit will also be energized. As mentioned previously, R20 is energized at the same time as the projector lamp. When R20 is activated the circuit paths to R21 and R22 are closed so that the continuity of the reward circuit is complete as shown in Figure 8. A correct choice will close the contacts of R21, thus shorting Contacts No. 1 and No. 2 of Timer No. 2 (reward delay) whereupon that timer begins operation (cf. Figure 4). The action of Timer No. 2 determines whether there will be a delay of reward and for how long; this is predetermined by *E* on the basis of the experimental design.

Timer No. 2 (reward delay) has been modified electrically to provide either normal operation or, alternatively, zero time delay. This feature allows *E* to preset the reward delay interval and then to use either zero delay or any other delay from 0.1 second to 111.0 seconds. This selection is accomplished by means of a single toggle switch in the Timer-Relay cabinet. If Timer No. 2 is set for immediate reward, it closes the contacts of Timer No. 1 (reward duration) without delay. If some finite interval is set into Timer No. 2, there will be a consequent delay in the completion of the impulse to Timer No. 1. Timer No. 1 controls the duration of the reward energizing both the R+ counter and the reward pen of the polygraph. The manual reward button on *E*'s control unit (cf. Figure 5) duplicates the function of R21; it will energize the reward light, the R+ counter, and the R+ pen.

The two boxes in Figure 4 labeled Clocks No. 1 and No. 2 denote the calibration instruments for reward and stimulus events. The connections shown supply power and signals for the Standard Electric clocks. Switch S28 in the Timer-Relay cabinet determines whether the first clock will monitor Timer No. 3 or Timer No. 4 and provides for accurate calibration of stimulus duration and dark period duration. Switch S29 determines whether the second clock will monitor Timer No. 1 or Timer No. 2 and provides calibration of reward duration and reward delay.

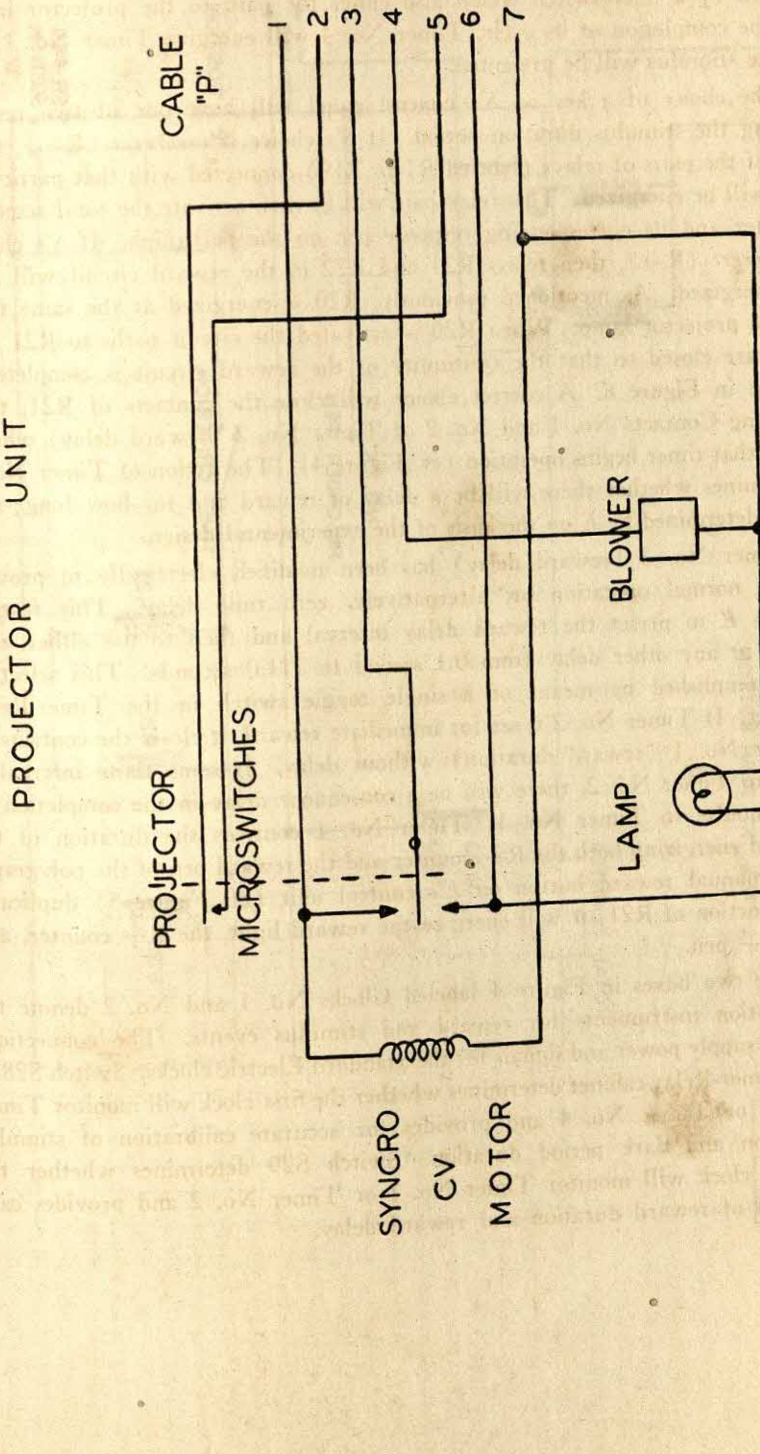


FIGURE 7
DIAGRAM OF PROJECTOR UNIT

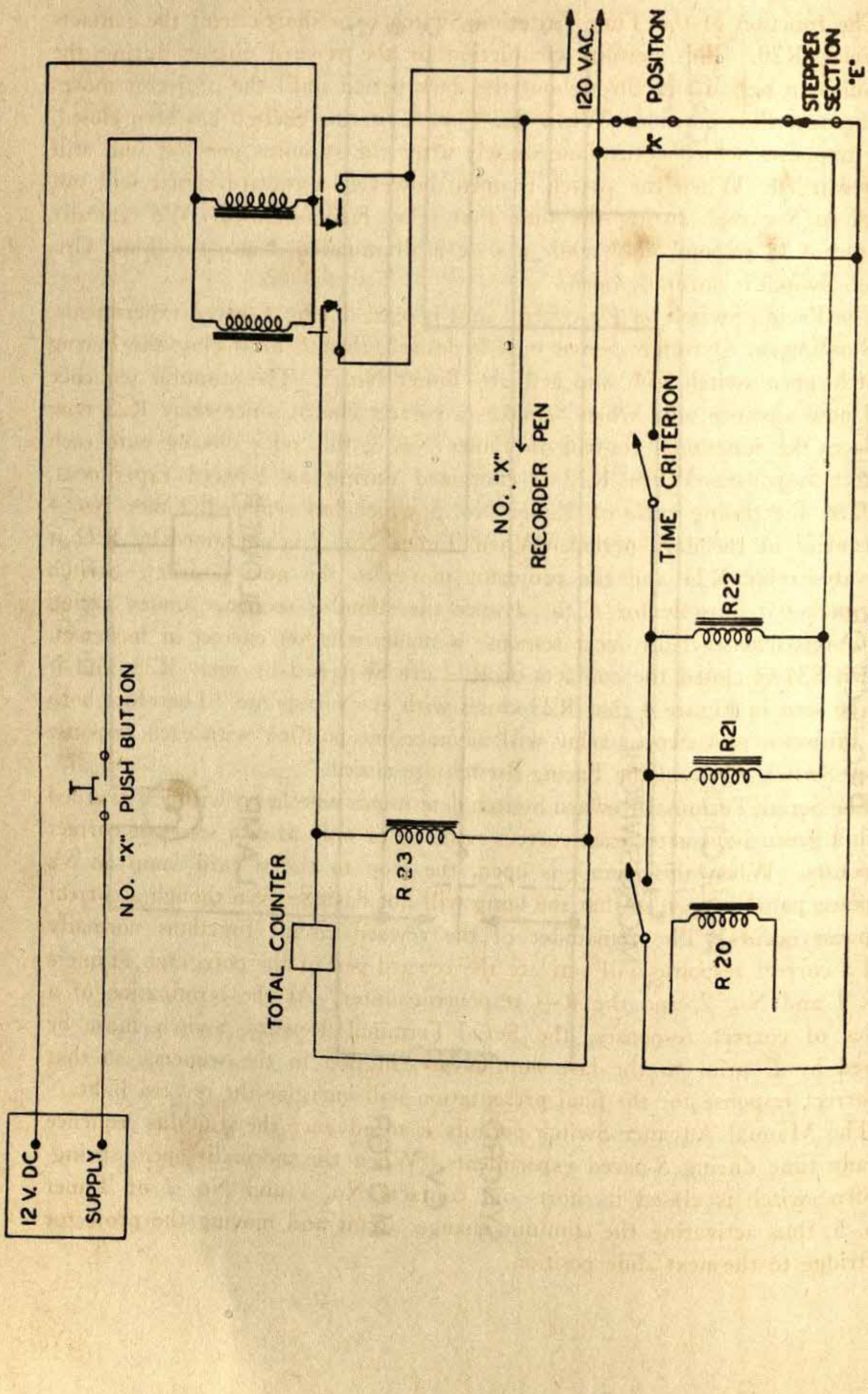


FIGURE 8
DIAGRAM OF REWARD CIRCUIT

The function of the Time Criterion Switch is to short-circuit the contacts of relay R20. This permits completion of the reward circuit during the stimulus-on period and throughout the dark period until the projector moves to the next slide position. When the Time Criterion Switch has been closed, late responses which occur immediately after the stimulus goes off will still be rewarded. When the switch is open, however, a reward signal will not be given *S* except during the time that relay R20 is closed. We typically instruct *S* to respond only while a slide is illuminated, hence the Time Criterion Switch is normally open.

The Pacing Switch on *E*'s control unit is open during *E*-paced experiments. If *S*-pacing on correct responses only is desired, then *E* must close the Pacing Switch, open switch S31, and activate Timer No. 5. The stimulus sequence will now advance only when *S* makes a correct choice, since relay R22 now replaces the functional control of Timer No. 3, this relay closing with each correct response. When R22 is energized during an *S*-paced experiment, it starts the timing cycle of Timer No. 5 which has replaced Timer No. 4 in control of the dark period. When Timer No. 5 is energized by R22 it activates relay R24 and the projector moves to the next position. Switch S31 makes it possible for *E* to advance the stimulus sequence under pacing conditions whenever an overt response is made, whether correct or incorrect. When S31 is closed the contacts of R22 are by-passed by relay R23, and it can be seen in Figure 4 that R23 closes with every response. Therefore both the projector and stepper relay will advance one position with each response when Switch S31 and the Pacing Switch are closed.

The Serial-Terminal Reward Switch determines whether *S* will be rewarded with a green light after each correct response or only after a series of correct responses. When this switch is open, the loop to the reward lamp on *S*'s response panel is open so that the lamp will not flash on even though a correct response occurs. The remainder of the reward circuit functions normally and a correct response will activate the reward pen of the polygraph, Timers No. 1 and No. 2, and the R+ response counter. At the termination of a series of correct responses, the Serial-Terminal Reward Switch must be closed by *E* prior to the last simulus presentation in the sequence so that a correct response for the final presentation will energize the reward light.

The Manual Advance Switch permits *E* to advance the stimulus sequence at any time during *S*-paced experiments. When this normally-open, spring-return switch is closed it shorts out contacts No. 1 and No. 2 of Timer No. 5, thus activating the stimulus change circuit and moving the projector cartridge to the next slide position.

The number of serial positions (slides) in a sequence is altered by changing the position of switch S30 in the Timer-Relay cabinet. A sequence may consist of any number of presentations up to 18, because recorder pen No. 2 is used for the reward signal. Switch S30 must always be set to include at least one more position than the number of stimuli in the sequence to allow for the homing action of the stepper when it resets to the first stimulus position. For example, if a trial consists of a sequence of 10 stimuli, then S30 must be set to position No. 11. This will cause the stepper to return to the first stimulus position after the 10th stimulus. The resetting of the stepper relay from the last stimulus position in a trial to the first stimulus position of the next trial is accomplished in the following manner. Relays R25A and R25B control the resetting of the stepper relay. Switch S30 is set to home the stepper following the end of the stimulus sequence. When this position is reached by the stepper, 120 v. AC power is applied through S30 and Section A of the stepper to R25A (cf. Figure 4). Relay R25A is closed and mechanically locked in position by R25B, where it will be held until released by R25B. When current flows through the closed contacts of R25A and Section B of the stepper relay to the stepper solenoid, the stepping switch advances continuously until position No. 23 is reached. At this point, A.C. power closes relay R25B through Section C and Contact No. 23, thus releasing the mechanical hold on R25A. The solenoid now automatically returns the stepper to the first position.

The length of a trial (number of choice points) is usually varied by increasing or decreasing the number of blank slides at the end of each series, but it can also be altered by changing the settings on Timer No. 3 or Timer No. 4 at the conclusion of a trial.

D. SUMMARY

In this paper the development of an automated circuit for the Selective Mathometer has been described, together with a portion of the new facilities of the Montana State University Psychology Laboratory. The Mathometer is an electro-mechanical pushbutton apparatus of analytic design and rugged construction for studying human multiple-choice learning phenomena. Compared with its semiautomatic predecessor of 1952, Mark II is an advanced system permitting automatic control of all temporal events, reward functions, and pacing variations. Achieving what we believe to be a pragmatic balance between standardized control and experimental versatility, the Selective Mathometer now provides the capability of investigating within a single reference task a wide range of acquired behaviors with relatively trouble-free operation and high statistical reliability.

REFERENCES

1. BILODEAU, E. A., & BILODEAU, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243-280.
2. NOBLE, C. E. Human trial-and-error learning. *Psychol. Rep.*, 1957, **3**, 377-398 (Monogr. Suppl. 8).
3. NOBLE, C. E., & FARESE, F. J. An apparatus for research in human selective learning. *J. of Psychol.*, 1955, **39**, 475-484.
4. NOBLE, C. E., NOBLE, J. L., & ALCOCK, W. T. Prediction of individual differences in human trial-and-error learning. *Percept. Mot. Skills*, 1958, **8**, 151-172 (Monogr. Suppl. 2).

Department of Psychology
Montana State University
Missoula, Montana

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 259-261.

TRANSITIONAL PROBABILITY IN ANAGRAM SOLUTION IN A GROUP SETTING*

Department of Psychology, Michigan State University

THOMAS STACHNIK[1]

A. PURPOSE

Mayzner and Tresselt have investigated the effect of different transition probabilities on anagram solution time (1). They explain transition probability as follows:

> In any given anagram or letter array there is a transition probability or digram frequency associated with each sequential pair of letters, and digram frequency tables are available which permit the relative assignment of a numerical value to this frequency. If the digram frequencies for all sequential pairs of letters of a given anagram are summed, one obtains a transition probability total for the anagram, which can be made to vary by permuting the letters of the anagram. A given letter set than can be arranged to form an anagram with a high transition total or a low transition total (p. 117).

The authors hypothesized that if a given anagram is presented in two forms, one in which the digram frequency total has been maximized, and one in which it has been minimized, the latter will be easier to solve since *S* will have less difficulty reorganizing the letters, i.e., the letter array will be easier to break up. Their results supported the hypothesis—the median solution time was significantly lower for anagrams with minimal transition totals when compared with the same anagrams with the letters arranged in maximal digram frequency totals.

The present study is an attempt to duplicate the Mayzner and Tresselt findings in a group setting. In the original study the anagrams were presented one at a time on index cards to individual *S*s, the dependent variable being the time required for solution. In the present study all anagrams were presented on a single sheet of paper, and each *S* was supplied with a pencil, the dependent variable being the number of correct solutions.

---

* Received in the Editorial Office on October 10, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author gratefully acknowledges the help given him by Dr. Donald M. Johnson.

B. METHOD

The students in two sections of introductory psychology at Michigan State University took part in the experiment ($N = 88$). The 20 anagrams employed were those used in the Mayzner and Tresselt study (1, Table 1). In order to control for individual differences in anagram-solving ability, the 40 anagrams (two different letter arrangements for each of the 20 correct solutions) were divided into two forms, each containing 10 "easy" anagrams (minimum digram frequency totals) and 10 "hard" anagrams (maximum digram frequency totals). The two forms were equated on the basis of median solution time in the Mayzner and Tresselt study, Form A anagrams requiring a mean solution time of 117.9 seconds and Form B 123.5 seconds (non-significant).

Each *S* received a single sheet of paper containing the 20 anagrams, roughly half of each group receiving Form A and the other half Form B. *E* explained that an anagram is a word whose letters are scrambled and that the task is to unscramble them and make a single five-letter word. *S*s were told that they would have 20 minutes to solve the 20 anagrams and were encouraged to try all 20 rather than spending too much time on any one since some of the anagrams were more difficult than others. At the end of 20 minutes the papers were collected and the answers given verbally by *E*.

C. RESULTS AND DISCUSSION

Available for analysis were two scores for each *S*, i.e., the number of "easy" and "hard" anagrams correctly solved, each score having a possible range of 0–10. The mean number of high transitional probability (TP) anagrams solved was 3.34 with a standard deviation of 2.30 and a range of 0–9. The mean number of low TP anagrams solved was 3.22 with a standard deviation of 2.31 and a range of 0–10. A paired-t test yielded a t of .22, so the difference between the means is within chance expectancy.

The negative results emphasize the differences between the individual administration of the Mayzner and Tresselt study and the group administration of the present study. In the former study the dependent variable was time required for solution, and the anagrams were presented one at a time on index cards to individual *S*s without paper or pencil. Since in the present study the anagrams were presented to *S*s with paper and pencil, it was possible to make overt trial-and-error attempts at solution, i.e., rearranging the letters on the paper for visual inspection. (Seventy-four of the 88 *S*s, or 84 per cent, actually did rearrange the letters on the paper as compared to only 14 *S*s who made only covert attempts at solution and wrote down the final answer.) Given a situation in which solution of a high TP anagram is attempted without paper and pencil, *S* covertly rearranges the letters, and when a new arrangement proves unsuccessful, he returns to the original letter arrangement and tries a new combination. In effect, then, *S* continues to work with an anagram whose TP remains constant. But if allowed to use paper and pencil, *S* tries new arrangements of the letters and the *TP* changes; if the new arrangement is unsuccessful, *S goes on from there* rather than returning to the original letter arrangement which represents the maximum digram frequency total for those particular letters. By rearranging the letters on paper and working from each subsequent revision, the TP is lowered, the letters become easier to reorganize, and hence the probability of solution is increased. Although this seems a reasonable explanation for the absence of a difference in the present study, further variations of the Mayzner and Tresselt procedure in both individual and group settings are necessary.

D. SUMMARY

An attempt was made to duplicate, in a group setting, the findings of an earlier study in which anagram solution time varied with the transitional probability of the anagram letter order, anagrams of high transitional probability yielding significantly longer solution times. The present study, like the previous study, utilized two levels of transitional probability by maximizing and minimizing digram frequency totals of 20 anagrams. Each *S* attempted to solve 10 anagrams of high transitional probability and 10 of low transitional probability, and a paired-*t* test indicated no difference in the number of anagrams correctly solved. A possible explanation was suggested as to why transitional probability is not a critical variable in anagram solution when administered to a group of *S*s using paper and pencil.

REFERENCE

1. MAYZNER, M. S., & TRESSELT, M. E. Anagram solution times: A function of transition probabilities. *J. of Psychol.*, 1959, **47**, 117-125.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 263-277.

PRIOR PAROCHIAL VS. SECULAR SECONDARY EDUCATION AND UTTERANCE LATENCIES TO TABOO WORDS*[1]

American International College

GEORGE S. GROSSER AND WILLIAM J. LACZEK

A. INTRODUCTION

The term "utterance latency" refers to the second of two portions of the word-recognition interval (that starting at word-recognition and ending at the correct enunciation of the word). "Perceptual defense" has been operationally defined as the excess of time required for the recognition of socially taboo words (or other emotional stimuli) in comparison with neutral words equated with the emotional words for length and for Thorndike-Lorge frequency count; see, e.g., definition (b) on page three of W. P. Brown's monograph (3). Such duration-thresholds of taboo-word recognition have been analyzed by J. S. Brown (2, pp. 303-304) as consisting of one or both of two possible components: (*a*) an emotional reaction distorting perception ("pure" perceptual defense); and/or (*b*) an emotional reaction resulting from correct perception but inhibiting the correct utterance of the taboo word presented.

Howes and Solomon (8), Postman, Bronson and Gropper (6), and J. S. Brown (2, pp. 316-319) have argued that so-called "perceptual defense" consists entirely of the second component above. W. P. Brown, however, does not reduce perceptual defense to utterance latency; instead, he proposes that words be scored on emotionality for each *S* by means of utterance latencies. Then the emotionality scores thus obtained would constitute an independent variable for recognition threshold measurements (see W. P. Brown, 3, pp. 29-44). He reports a curvilinear relation between stimulus emotionality and recognition threshold.

Utterance latencies (the timing of which begins when the stimulus is turned off) may be affected by one of two components: (*a*) The utterance latency measure is equivalent to a disjunctive reaction time, in which the

---

* Received in the Editorial Office on March 19, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was initiated as an experiment in the course in Experimental Psychology at American International College, Springfield, Massachusetts. The course instructor was the senior author while the two students who originated the study were the junior author and Mr. Thomas P. Leyden.

number of disjoined stimuli is the size of the *S*'s vocabulary. If any set is introduced, the number of disjoined stimuli is reduced; if *S* becomes set for, e.g., socially taboo words, the number of possible stimuli is reduced to the size of the *S*'s taboo-word subvocabulary. (*b*) Response suppression may still operate to raise the *S*'s latency of utterance.

The advantage of the use of the utterance latency rather than the duration-threshold as a dependent variable lies in the possibility of parcelling out the factors of set and of response suppression since the two exert opposed influences on the utterance latency. (A disadvantage is the lack of immediate generalizability of utterance latency data to the great body of duration-threshold data.)

Postman, Bronson and Gropper (6) have shown that women had higher recognition thresholds than men when taboo words were presented to them. The present experiment involves in part a further test of this hypothesis. Moreover, the variable of parochial vs. secular schooling may have "naturally" varied the emotional strength of responses aroused by taboo words; it was hypothesized that parochial secondary school graduates would take longer to utter taboo words than secular secondary school graduates.

To parcel out set-change from response-suppression, (*a*) a set of aggressive words was interposed between a set of neutral and a set of taboo words; and (*b*) a second set of neutral words followed the set of taboo words. This would allow the demonstration of any increase in latency attributable to set change alone. All words were to be presented tachistoscopically to graduates of parochial and of secular secondary schools. They were to enunciate each word as soon as it was recognized.

B. METHOD

1. *Subjects*

The *S*s were selected from the freshman class at American International College, Springfield, Massachusetts. There were two groups: (*a*) students from parochial secondary schools, and (*b*) students from secular secondary schools. There was no knowledge on the part of the experimenters of the group membership of any *S*. Twenty-four *S*s were used, 12 parochial and 12 non-parochial. In each group, there were five females and seven males.

2. *Apparatus*

A Keystone "Flash Meter" projector, equipped with a shutter, was used to project words (typed in capitals) upon a portable screen six feet from the *S*. A Springfield timer was operated by the *E* to register the *S*'s utterance latency. A second *E* (the recorder) entered the raw data in a data sheet.

3. *Procedure*

Each *S* was brought into the experimental room and seated to the right of the tachistoscope and several inches behind it. All slide materials and raw data sheets were out of the *S*'s sight. The *E* said: "I want you to concentrate on the screen and tell me what you see." The *E* then set the time exposure at its minimum value and exposed the stimulus word by depressing a lever on the shutter. After the *S* reported the stimulus word, the *E* continued to establish the point of visual recognition. At the lower limen each *S* was able to perceive each stimulus word. His task was to recognize and to report each word.

Each *S* was given several adaptation trials with neutral words not to be used in the experimental presentations. Examples of these are: verify, income, commerce, and elegant. These words, along with the thirty neutral experimental words, were all equated on the Allport-Vernon Study of Values (1) according to the six categorical values: theoretical, economic, aesthetic, political, social, and religious. All words except the taboo words were equated as to word frequency by the Lorge-Thorndike Word Count (7). The thirty neutral words were all of equal familiarity based on frequency of occurrence. The fifteen aggressive words were also equated in the same manner; all the words fell in the range of 28 to 52 occurrences of each word per million taken from the "T" counts, which indicate frequency in readers, text books, the Bible, and the English classics. After *S*s were familiarized with the screen, the tachistoscope, and the exposure time, they were instructed: "Remember, as soon as you see or recognize the word, I want you to tell me what the word is. For instance, if the word that appears on the screen is DOG, then immediately say 'dog.' All right, are we ready to begin?" Before each stimulus word, the *E* checked with the recorder to verify that the response had been recorded. He then said "Ready," alerting the *S* for the word to follow.The timer yielded the latencies as follows: the stimulus word was flashed on the screen at the *S*'s previously determined lower recognition limen. When the word disappeared, the recorder depressed a cord switch with his thumb, starting the timer; at the instant *S* made an oral response, the recorder stopped the timer and the reading was taken. The time units were in hundredths of a second with one complete rotation of the dial being one full second. The order of presentation was the same for each *S*: 15 neutral words, 15 aggressive words, 15 sex words, and the last 15 neutral

TABLE 1
EXPERIMENTAL WORD SERIES WITH THORNDIKE-LORGE *T* RATINGS

Neutral	T Count	Aggressive	T Count	Taboo	T Count	Neutral	T Count
Analysis	50	Strangle	57	Prostitute	16	Loving	?
Research	28	Slit	28	Sperm	28	Kindly	310
Graceful	97	Smash	28	Homosexual	none	Devoted	30
Wealthy	160	Mangle	50	Pervert	50	Artist	160
Logical	12	Slug	57	Adultery	18	Blessed	?
Helpful	90	Slam	57	Douche	none	Worship	235
Friendly	320	Deface	28	Intercourse	57	Sacred	210
Literary	90	Hack	57	Erection	18	Reverent	25
Science	200	Slap	57	Lesbian	none	Famous	700
Prayer	350	Brutal	47	Seduce	57	Dominate	18
Complete	50	Annihilate	28	Vagina	none	Politics	28
Citizen	260	Maim	57	Penis	none	Govern	240
Religion	210	Smear	50	Masturbation	none	Poetry	130
Finance	57	Maul	28	Rape	28	Sociable	12
Economic	28	Slash	50	Incest	none	Beauty	700

words. The stimulus words, in the actual order of presentation, are presented in Table 1 along with their respective Thorndike-Lorge frequencies.

The *S* was allowed five trials for each word; however, no two such trials were consecutive, and the repetition of any stimulus word was done in a semi-random order. The only stipulation was that three consecutive stimulus words be correctly reported before the repetition was presented during the period in which that type of word occurred. For example, an aggressive or taboo word would not be shown among neutral words, and *vice versa*. Both the *E* and the recorder varied procedure somewhat in order to counterbalance any cues that the *S* might employ. The only cue used was the "Ready" signal given before presentation of the stimulus. The *E* lowered the time of exposure on a few occasions (as "catch" trials) in order to break any experimental set and then reset it to its original value before continuing with the next word.

For each *S*, a total of sixty reaction times were taken, one for each stimulus word. The RT was a measure of the time it took from the termination of the stimulus to the onset of the oral response. The RTs were transformed into their reciprocals.

C. Results

Table 2 includes the mean reciprocal RTs to each of the four sets of stimulus words. For every *S*, the smallest reciprocal is that of the mean RT to the taboo words.

Error terms for use in the analysis of variance described below were subjected to Bartlett's test of homogeneity of variance. The variances within the four experimental groups (male secular, female secular, male parochial, female parochial) could be pooled into a single error term, since chi-square was 5.74 and the null hypothesis was accepted. The four variances of the groups × word-blocks interaction also could be pooled into a single error term, since chi-square was 4.99 and the null hypothesis was accepted.

An analysis of variance was done to test the following main effects: parochial vs. non-parochial secondary schooling; sex of *S*; and word-set (first neutral, I; aggressive, II; taboo, III; second neutral, IV), as well as the following interactions: education × sex; sex × word-blocks; education × word-blocks; and sex × education × blocks. This analysis is presented in Table 3.

The main effect for blocks and the education × blocks interaction proved significant at the .01 and .05 levels, respectively.

A further analysis of the blocks main effect was done using Scheffé's test

TABLE 2
MEAN RECIPROCALS OF REACTION TIMES TO EACH OF THE FOUR SETS OF WORDS
The Reciprocal Scale is in 1/Second Units

Subject	First neutral	Aggressive	Taboo	Last neutral
FP-1	1.375	1.167	.702	.965
FP-2	1.198	1.082	.644	1.079
FP-3	1.094	.974	.760	1.656
FP-4	1.053	1.018	.836	1.282
FP-5	1.116	1.098	.668	1.172
MP-6	1.117	.934	.728	.920
MP-7	1.254	1.344	.833	1.140
MP-8	1.184	.980	.728	1.056
MP-9	1.283	1.370	.894	1.500
MP-10	.738	.729	.589	.740
MP-11	1.039	.992	.617	.991
MP-12	1.325	1.215	1.059	1.404
FNP-1	.952	.958	.944	.955
FNP-2	1.286	1.098	1.032	1.237
FNP-3	1.125	.989	.905	1.167
FNP-4	1.073	1.042	.648	1.092
FNP-5	1.268	1.213	.957	1.203
MNP-6	1.145	1.082	.810	1.032
MNP-7	1.012	.915	.746	1.007
MNP-8	.937	.882	.758	.805
MNP-9	.788	.800	.629	.820
MNP-10	1.176	1.106	.892	1.094
MNP-11	1.256	1.039	.834	1.119
MNP-12	1.278	.938	.779	1.152

TABLE 3
ANALYSIS OF VARIANCE OF RECIPROCAL RTs

Sources of variation	df	SS	Mean S	F	P
Between subjects	23	1.698			
Sex	1	.088	.088	1.14	NS
Education	1	.028	.028	—	NS
Sex × education int.	1	.037	.037	—	NS
Between *S*s within grps.	20	1.546	.077		
Within subjects	72	2.584			
Blocks	3	1.736	.579	50.77	.01
Sex × blocks int.	3	.036	.012	1.06	NS
Education × blocks int.	3	.102	.034	2.97	.05
Sex × educ. × block int.	3	.026	.009	—	NS
Pooled *S*s × blk × within Sex-education group int.	60	.684	.011		
Total	95	4.281			

for multiple comparisons (4, 154-156). The orthogonal comparisons that were made appear in Table 4.

The only significant comparisons involve the taboo words (*a*) against any one of the other blocks of words; (*b*) combined with the aggressive

TABLE 4

ORTHOGONAL COMPARISONS, USING THE SCHEFFÉ METHOD, OF SUMS OF RECIPROCAL REACTION TIMES FOR EACH WORD BLOCK

$$A_i = \frac{D_i^2}{n\Sigma a_i^2} \qquad n = 24$$

Block I: First neutral
Block II: Aggressive
Block III: Taboo
Block IV: Second neutral

Comparison		Block I	II	III	IV	D	D^2	A
I vs. IV	a	+ 1	0	0	− 1			
	aΣX	+27.072	0	0	−26.628	6.444	0.971	0.004
I vs. II	a	+ 1	− 1	0	0			
	aΣX	+27.072	−24.965	0	0	2.107	4.439	0.092
I vs. III	a	+ 1	0	− 1	0			
	aΣX	+27.072	0	−18.992	0	8.080	65.286	1.360*
II vs. III	a	0	+ 1	− 1	0			
	aΣX	0	+24.965	−18.992	0	5.973	35.677	0.743*
I and IV vs. II and III	a	+ 1	− 1	− 1	+ 1			
	aΣX	+27.072	−24.965	−18.992	+26.628	9.743	94.926	0.989*
I, II, and IV vs. III	a	− 1	− 1	+ 3	− 1			
	aΣX	−27.072	−24.965	+56.976	−26.628	21.689	470.413	1.633*

* Significant at the .05 level.

words as against the neutral words; and (*c*) against all three other blocks of words.

The reciprocal reaction-times for taboo words are significantly lower in each comparison; i.e., correct recognition responses are given more slowly to these stimuli than to any of the other sets of words.

A *t*-test was done between the mean reciprocal RT of the parochial group for the taboo words and the mean reciprocal RT of the non-parochial group for that block. A *t* of 1.35 was obtained; for 22 degrees of freedom, this value of *t* is not significant.

A *t*-test was also done between the mean reciprocal RT of the parochial group for the second block of neutral words and the mean reciprocal RT of the non-parochial group for that block. A *t* of 13.74 was obtained; for 22 degrees of freedom, this is significant at the .005 level. That is, the parochial *S*s as a group react more quickly to these neutral words than the non-parochial *S*s do.

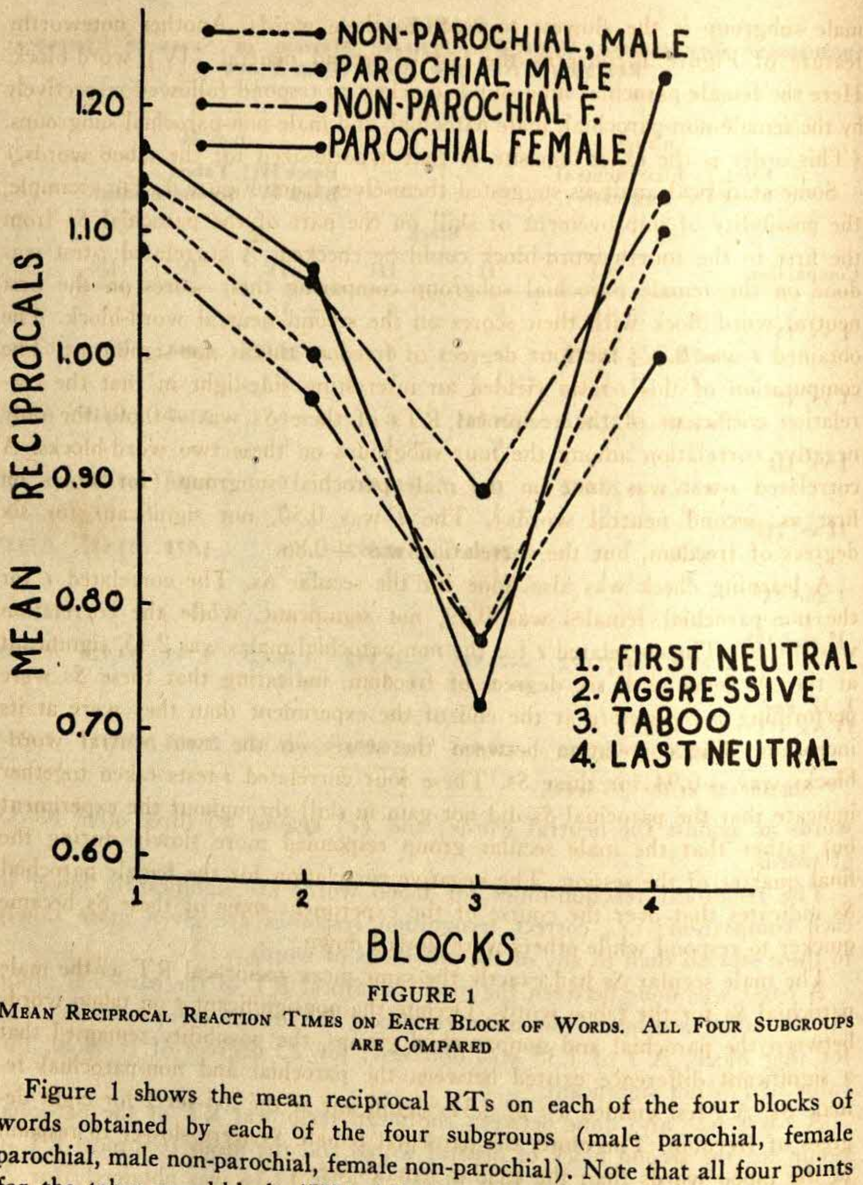


FIGURE 1

MEAN RECIPROCAL REACTION TIMES ON EACH BLOCK OF WORDS. ALL FOUR SUBGROUPS ARE COMPARED

Figure 1 shows the mean reciprocal RTs on each of the four blocks of words obtained by each of the four subgroups (male parochial, female parochial, male non-parochial, female non-parochial). Note that all four points for the taboo word-block (III) fall below the lowest of the other points. It is also evident that the non-parochial female subgroup has the highest speed of reaction (reciprocal RT) to taboo words while the parochial female subgroup is the slowest to react to these words. Another noteworthy feature of Figure 1 concerns the last or second neutral (IV) word-block. Here the female parochial subgroup is quickest to respond followed respectively by the female non-parochial, male parochial, and male non-parochial subgroups. (This order is the exact opposite of that hypothesized for the taboo words.)

Some statistical analyses suggested themselves from Figure 1. For example, the possibility of improvement of skill on the part of the parochial *S*s from the first to the fourth word-block could be checked. A correlated *t*-test was done on the female parochial subgroup comparing their scores on the first neutral word-block with their scores on the second neutral word-block. The obtained *t* was 0.72; for four degrees of freedom this is not significant. The computation of this *t*-ratio yielded an interesting side-light in that the correlation coefficient of the reciprocal RTs of these *S*s was —0.69, the only negative correlation among the four subgroups on these two word-blocks. A correlated *t*-test was done on the male parochial subgroup (for scores on first vs. second neutral words). The *t* was 0.50, not significant for six degrees of freedom, but the correlation was +0.86.

A learning check was also done for the secular *S*s. The correlated *t* for the non-parochial females was 0.09, not significant, while the correlation was +0.97. The correlated *t* for the non-parochial males was 2.45, significant at the .05 level with six degrees of freedom, indicating that these *S*s were performing *more slowly* at the end of the experiment than they were at its inception. The correlation between the scores on the two neutral word-blocks was +0.94 for these *S*s. These four correlated *t*-tests taken together indicate that the parochial *S*s did not gain in skill throughout the experiment but rather that the male secular group responded more slowly during the final quarter of the session. The negative correlation for the female parochial *S*s indicates that over the course of the experiment some of these *S*s became quicker to respond while others were slowed down.

The male secular *S*s had exactly the same mean reciprocal RT as the male parochial *S*s for the taboo words. Despite the nonsignificant *t* on taboo words between the parochial and nonparochial groups, the possibility remained that a significant difference existed between the parochial and non-parochial female *S*s. A *t*-test yielded a *t* of 2.31, significant at the .05 level for eight degrees of freedom. Therefore, the parochial female *S*s were slower to respond to the taboo words than the secular females were.

Figure 2 compares the parochial and non-parochial groups on the four blocks of words, omitting the sex-of-*S*s variable. Some *t*-tests were done that bear on this graph.

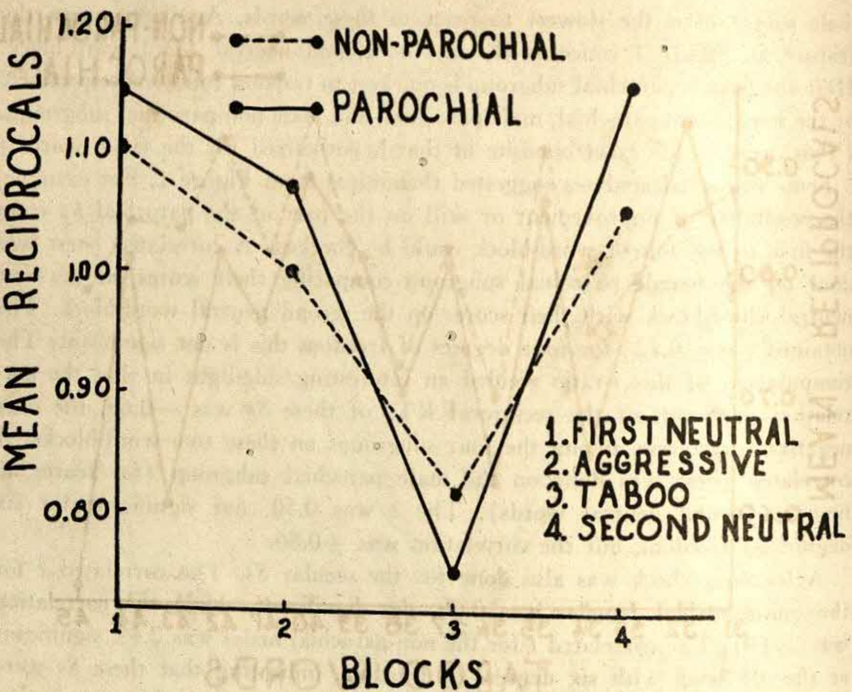


FIGURE 2
MEAN RECIPROCAL REACTION TIMES ON EACH BLOCK OF WORDS. PAROCHIAL AND NON-PAROCHIAL DATA ARE COMPARED

The correlated *t* comparing scores on first and last neutral words for the secular *S*s was 2.30, significant at the .05 level with eleven degrees of freedom; the correlation coefficient was +0.89. The performance of the *S*s was significantly slowed over the course of the experiment; however, this result is attributable to the male *S*s (see above).

The correlated *t* comparing scores on the first and last neutral words for the parochial *S*s was 0.16, not significant with eleven degrees of freedom; the correlation coefficient was +0.40. The low correlation is due to a negative correlation for the female *S*s (see above).

The *t*-ratio comparing these two groups on the second word-block (aggressive words) was 0.33, not significant with 22 degrees of freedom. Above we have indicated that a like comparison on taboo words yielded a non-significant *t*, although the female *S*s did differ (according to type of education) on taboo words.

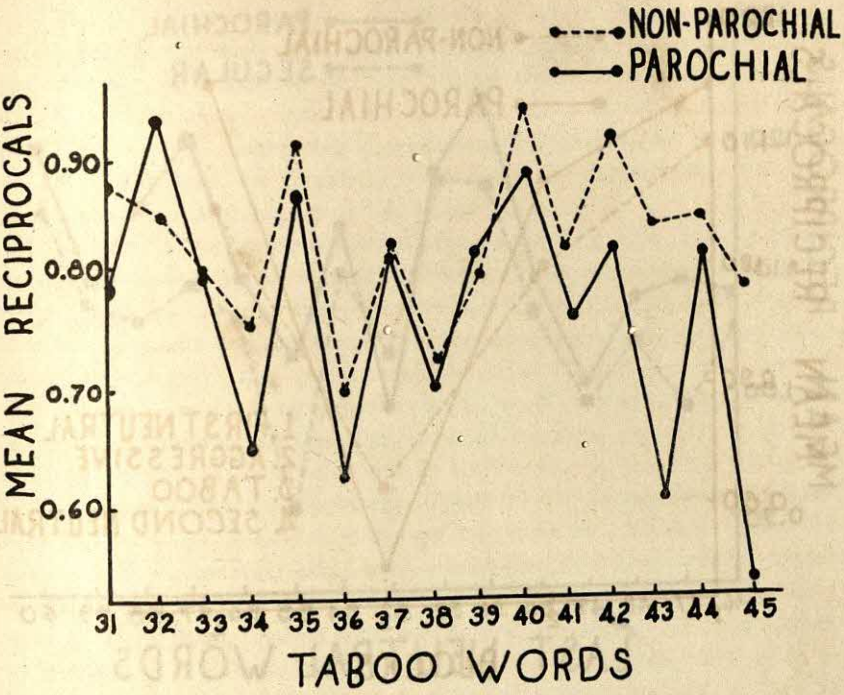


FIGURE 3

Mean Reciprocal Reaction Times on Each Word of Taboo Series. Non-parochial and Parochial Data are Compared

Figure 3 indicates the mean reciprocal RTs of the two groups on the individual words comprising the block of taboo words. The graph shows that the groups seem to differ especially over the last six words (words 40 through 45). The *t*-ratio on these six words was 2.28, significant at the .05 level with 22 degrees of freedom. The secular *S*s responded more quickly than the parochial *S*s to these six words.

Figure 4 shows the mean reciprocal RTs of the two groups on the individual items of the last neutral word-block.

D. Discussion

Among the intervening variables that have been postulated to account for the typical perceptual defense data are set (6) and response-suppression (5). The present results can be interpreted so as to parcel out the effects of these two constructs. Let us designate the difference in mean reciprocal RT (i.e., in speed of response) between the first neutral words and the taboo list as

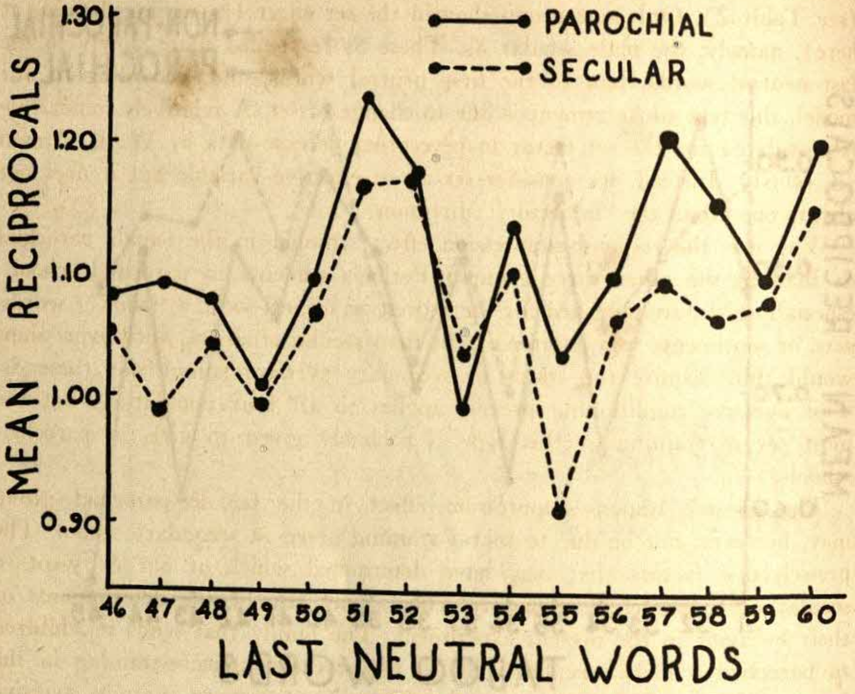


FIGURE 4
MEAN RECIPROCAL TIMES ON EACH WORD OF LAST NEUTRAL SERIES. NON-PAROCHIAL AND PAROCHIAL DATA ARE COMPARED

(IN–IIIT). This quantity should reflect both effects, so the (set) + (response-suppression) = (IN–IIIT). Low response-speeds could be due to a sudden change in set from the "neutral" or "aggressive" to the "taboo" category of words. These low speeds may also result from response-suppression.

The change from the taboo words (IIIT) back to the final set of neutral words (IVN) should lead to recovery from the effects of response-suppression but to a new loss of response-speed due to a further change of set (especially since the number of disjoined stimuli is raised to the full size of the *S*'s vocabulary). Hence any recovery of response-speeds to the IVN words would be a measure of the effect of response-suppression due to the IIIT words: response-suppression effect = (IVN–IIIT). This also enables us to estimate the effect of set: (IN–IIIT)—(IVN–IIIT) = (IN–IVN) = the increase in speed of response attributable to set.

The results indicate that all four groups exhibited the effects of response-suppression, although this is most pronounced for the parochial female *S*s (see Table 2). Only one group showed the set effect (as we are defining it here), namely, the male secular *S*s. These *S*s responded more slowly to the last neutral words than to the first neutral words, and, according to our model, this type of decrement is due to change of set. A relatively minor role is postulated for the set factor in perceptual defense data by W. P. Brown (3, 60-61). Instead, we consider set as an effective variable but it does not always operate in the "inhibitory" direction.

Why was the response-suppression effect stronger in the female parochial *S*s than in the other three groups? Perhaps students in parochial schools, especially girls, are punished for the expression of anti-social ("taboo") words, acts, or sentiments to a greater extent than secular students. Such expressions would thus acquire the status of secondary aversive stimuli for these *S*s. This aversive conditioning process applies to all four populations but the most severe training of this type is probably given to girls in parochial schools.

The greater response-suppression effect in the female parochial group may, however, not be due to moral training given in secondary school. The pre-selective factors that may have determined which of our *S*s went to secular or to parochial secondary schools may also be the determinants of their behavior in the present experiment. The family that sends its children to parochial schools may be likely to use severe conscience-training in the home. This would predispose such children to react to morally aversive situations by response suppression. If carried to a neurotic extreme these children would learn to prefer the defense mechanism of repression.

In the case of the secular groups, early home training or secondary education might stress intellectual values more than moral ones. If so such children or adolescents would learn to prefer the defense mechanism of intellectualization.

W. P. Brown (3) has analyzed perceptual defense data in personality terms and finds two trait-clusters, one typical of *S*s with high recognition thresholds (slow responders in the present study), the other representing *S*s with low thresholds (our fast responders). The former combine the following traits: recall more completed tasks than incomplete ones; high hysteria score on the MMPI; low manifest anxiety; and extraversion. The fast responders have the complementary traits: recall more incomplete than completed tasks; high manifest anxiety; high psychasthenia score on the MMPI; and introversion. Tentatively, we would classify our female parochial *S*s with the slow responders and our female non-parochial *S*s with the fast responders.

Another suggestion from the present study would be to use scores on the theoretical and religious scales of the Allport-Vernon Study of Values to predict which *S*s would be fast and which slow responders to taboo words.

One word in the aggressive (II) word-block seemed to give all the *S*s a little more trouble than any of the others, viz., *annihilate*. This would seem to be a question of familiarity, although its "T"-count frequency of occurrence was comparable to that of the other 14 aggressive words. In comparison with taboo words of equal length, e.g., "homosexual," "masturbation," "intercourse," and "prostitute," the aggressive word "annihilate" yielded a larger latency for most *S*s in either group. We question the applicability of the Lorge-Thorndike count to these five words, which are rated equal in frequency. The taboo words are better known to the freshmen we observed than is "annihilate."

E. SUMMARY

Two groups of 12 *S*s each were chosen from a college freshman class. One group consisted of graduates of parochial schools, while the other was made up of graduates of secular schools. Each group had seven men and five women. All *S*s were shown tachistoscopically four word lists—neutral (I), aggressive (II), sexually taboo (III), and neutral (IV). All lists sampled the same Lorge-Thorndike word-count range, although some taboo words were of lower "T" count than the remainder of the words. In addition, the neutral words each represented one of the six Allport-Vernon value scales. All *S*s responded more slowly to taboo words than to other words. The effect was the most pronounced for the parochial female *S*s and the least pronounced for the secular female *S*s. Due to the use of the four word lists and to the use of the utterance latency as the dependent variable, it was possible to set up a model separating the effects of change of set from those of response-suppression. Response-suppression accounted in whole or in part for the data from all *S*s; some of the data for male secular *S*s could be attributed to change of set.

REFERENCES

1. ALLPORT, G. W., & VERNON, P. E. A Study of Values. Boston: Houghton-Mifflin, 1931.
2. BROWN, J. S. The Motivation of Behavior. New York: McGraw-Hill, 1961.
3. BROWN, W. P. Conceptions of perceptual defence. *Brit. J. Psychol. Monog. Supplements*, 1961, **35**.
4. EDWARDS, A. L. Experimental Design in Psychological Research. (2nd ed.) New York: Rinehart, 1960.
5. HOWES, D. H., & SOLOMON, R. L. A note on McGinnies, "Emotionality and perceptual defense. *Psychol. Rev.*, 1950, **57**, 229-234.

6. Postman, L., Bronson, W. C., & Gropper, G. L. Is there a mechanism of perceptual defense? *J. Abn. & Soc. Psychol.*, 1953, **48**, 215-224.
7. Thorndike, E. L., & Lorge, I. The Teacher's Word Book of 30,000 Words. New York: Columbia, 1944.

American International College
Springfield 9, Massachusetts

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 279-286.

DEVELOPMENT OF PATRIOTISM IN CHILDREN— A SECOND LOOK*[1]

State University of New York at Albany

EDWIN D. LAWSON

A. INTRODUCTION

The study and evaluation of attitudes of patriotism or nationalism has been of concern to the social psychologist in the past and is of continuing concern. Investigations of this dimension of attitudes are relatively few with the exception of Horowitz (1) and Newcomb (2).

The Horowitz study was the first investigation of the development of attitudes of children at elementary school level. He reported the growth of patriotism in 82 children, Grades 1-10 in a rural four-room school in Tennessee. As his measure of patriotism, Horowitz had the children choose which of 24 flags were ranked in the top five on the basis of "best-looking." Lawson (3) and Lawson and Stagner (4) also used reactions to the beauty of flags as a measure of nationalism. College students' responses to a flag scale were shown to have a significant correlation with the Ferguson Nationalism Scale (5), a Thurstone-type attitude scale.

Horowitz in his study reported that:

1. Appreciation of the U. S. flag shows a steady development from 27.3 per cent in Grade 1 to 100 per cent from Grade 7 on. (Score was determined by the percentage of pupils who ranked the U. S. flag among the top five.)
2. The best-liked flag by the younger children was the Siamese Merchant flag which showed a white elephant.
3. Flags in some way similar to that of the U. S. (Liberia, China, Cuba) tended to be preferred at higher grade levels.

The purpose of this investigation was to determine with a larger sample:

1. Whether responses 25 years later would be similar; i.e., would there be a steady development of appreciation for the U. S. flag, positive

---

* Received in the Editorial Office on October 24, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Revision of a paper read at the annual meeting of the Eastern Psychological Association, April, 1962. This investigation was carried out with the collaboration of Robert J. Cornell, Elizabeth Cunningham, Mary C. Gatzemeyer, Elaine Heidenstrom, Helga Karker, Alvin Rabinowitz, John Ray, and Patricia Wolcott.

reaction to flags with animals, and a generalization pattern for flags similar to that of the U. S.?

2. What responses would be made to the Soviet flag and the United Nations flag.
3. Whether an emergent perceptual and choice pattern could be identified.

B. METHOD

1. *Subjects*

The *S*s were children from two school systems (one urban, Schenectady; one suburban, Averill Park) in upstate New York. Twenty boys and 20 girls, kindergarten through twelfth grade (K-12), were interviewed in each system. Thus a total of 1040 school children was tested.

2. *Procedure*

The procedure, while similar to that of Horowitz, did have some differences. Twenty flags were mounted separately on small masonite rectangles and were randomly presented in a matrix 5 by 4. Each flag was about 1-5/8 inches by 1-1/8 inches and was cut from a chart obtained from the United Nations Office of Public Information. The U. N. flag was slightly reduced in size in order to be equal to the others. *S*s were interviewed separately and were asked to pick out the most attractive flag, then the next most attractive flag, until all of the flags had been ranked and their position recorded by the interviewer. The Ferguson Nationalism Scale was also administered to Grades 10-12. The interviewers were eight graduate students in social psychology.

C. RESULTS

The ranks chosen for each flag were recorded by grade level and sex in each school system. Since the data from the two systems and of boys and girls showed a strong similarity, the data were combined. Following Horowitz (1), Table 1 shows the percentage of pupils, K-12, who included the respective countries among their first five choices. Figure 1 is a comparison of the development of appreciation of the flags of the U. S. and Liberia in this study (data obtained in 1961) with that of the Horowitz (data obtained in 1936, reported in 1940). Figure 2 shows the development of attitudes toward the U. S., the U. N., Liberian, and Soviet flags in this study.

1. *U. S. Flag*

The U. S. flag is rated highest by the 1961 school children at all levels except Grades 10 and 11, but the pattern is quite different from that of the

TABLE 1

The Percentage of Students in Each Grade, K-12, Who Ranked Flags in the Top Five*

	Grades K	1	2	3	4	5	6	7	8	9	10	11	12
United States	73.75	68.75	78.75	81.25	73.75	76.75	78.75	75.00	71.25	72.50	65.00	70.00	68.75
United Nations	25.00	12.50	18.75	15.00	31.25	36.25	52.50	66.25	50.00	60.00	68.75	82.50	60.00
Russia	10.00	6.25	7.50	7.50	2.50	5.00	5.00	0.00	5.00	3.75	3.75	5.00	1.25
Liberia	55.00	47.50	58.75	58.75	50.00	51.25	26.25	48.75	22.50	33.75	21.25	26.25	28.75
Greece	28.00	28.75	32.50	32.50	17.50	23.75	25.00	47.50	25.00	36.25	33.75	33.75	42.50
Israel	26.25	28.75	18.75	16.25	10.00	21.25	16.25	20.00	18.75	36.25	15.00	32.50	12.50
Turkey	26.25	15.00	17.50	10.00	8.75	11.25	12.50	8.75	8.75	11.25	5.00	8.75	7.50
Yemen	13.75	13.75	17.50	16.25	12.50	13.75	11.25	11.25	6.25	22.50	10.00	18.75	10.00
France	23.75	20.00	31.25	17.50	12.50	16.25	16.25	12.50	9.75	22.50	17.50	12.50	17.50
Norway	18.50	12.50	13.75	11.25	12.50	12.50	18.75	15.00	5.00	15.00	6.25	2.50	11.50
China	20.00	29.25	13.75	20.00	13.75	6.25	5.00	5.00	10.00	7.50	1.25	6.25	5.00
Great Britain	15.00	15.00	26.75	30.00	28.75	26.25	31.25	20.00	23.75	26.25	32.50	22.50	28.75
Saudi Arabia	20.00	21.25	20.00	21.25	15.00	16.25	21.25	22.25	22.25	18.75	18.75	18.75	22.50
Brazil	20.00	23.00	22.50	26.25	35.00	16.25	34.25	20.00	28.75	13.75	16.25	22.25	20.00
Guatemala	15.00	26.25	18.75	27.50	52.50	40.00	37.50	30.00	58.75	46.25	51.25	46.25	46.00
Peru	20.00	26.25	17.50	21.25	26.25	28.00	25.00	20.00	28.75	23.75	31.25	22.50	28.75
Iran	25.00	31.25	28.00	31.25	25.00	21.25	17.50	15.00	15.00	12.50	28.75	22.50	13.75
Canada	21.25	20.00	30.00	17.50	33.75	28.75	30.00	30.00	28.75	21.25	40.00	17.50	29.25
Lebanon	27.50	35.00	18.75	23.75	21.25	17.50	18.75	15.00	21.25	20.00	21.25	32.50	11.25
Byelorussia	15.00	16.25	8.75	16.25	13.75	28.75	13.75	10.00	16.25	8.75	20.00	10.00	18.75

* Thus 25 per cent of the children in kindergarten ranked the U.N. in the top five.

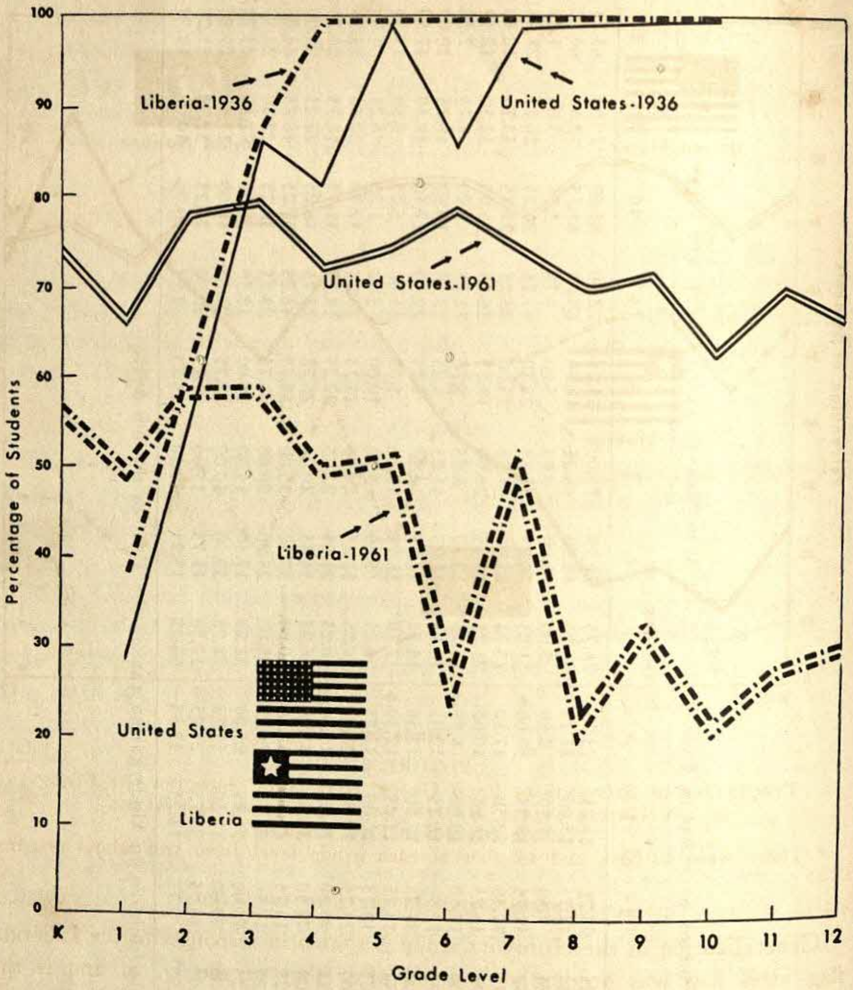


FIGURE 1

Development of Appreciation of Flags of the United States and Liberia*

* Expressed in percentage of students who ranked these in the first five.

Horowitz study as is shown in Figure 1. In our study, appreciation is fairly constant from kindergarten on with an average percentage of about 70 (indicating that 70 per cent of the children ranked the U. S. flag in the top five). In contrast Horowitz showed a positive acceleration from 27 per cent in Grade 1 to 100 per cent in Grade 7.

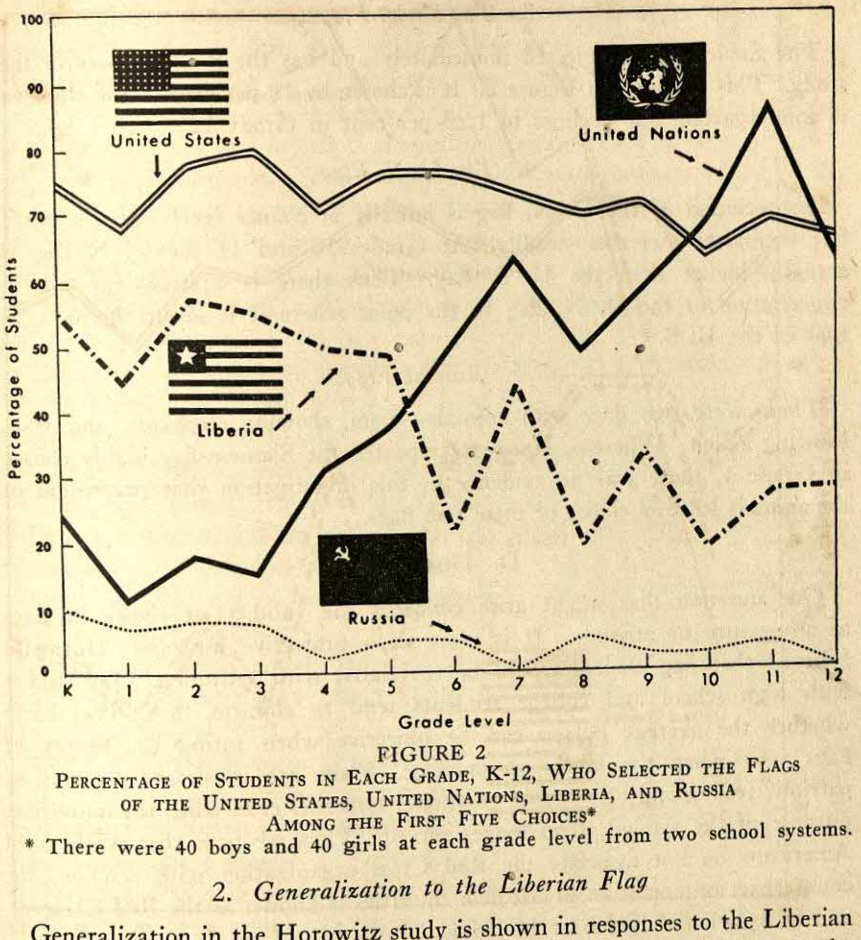


FIGURE 2

PERCENTAGE OF STUDENTS IN EACH GRADE, K-12, WHO SELECTED THE FLAGS OF THE UNITED STATES, UNITED NATIONS, LIBERIA, AND RUSSIA AMONG THE FIRST FIVE CHOICES*

* There were 40 boys and 40 girls at each grade level from two school systems.

2. *Generalization to the Liberian Flag*

Generalization in the Horowitz study is shown in responses to the Liberian flag (this flag was originally modeled after that of the U. S. and is the most similar). While the 1961 children did show an initial preference for the Liberian flag at the kindergarten level (55 per cent) which is substantially above a chance score of 25 per cent but below that of the U. S., this score did not increase or even maintain this level but dropped to a percentage just above chance in the high school group. While generalization does tend to occur somewhat in the case of Liberia, it is greatest at the elementary school level and then tends to diminish. There is a slight tendency to favor China in Grade 1 in the 1961 sample. This is another flag mentioned by Horowitz as being similar in pattern to that of the U. S.

3. *The Soviet Flag*

The Soviet flag is rejected immediately and has the lowest scores in the study. This is shown in Figure 2. It is chosen by 10 per cent of the children in kindergarten, but declines to 1.25 per cent in Grade 12.

4. *The U. N. Flag*

Appreciation of the U. N. flag is initially at chance level (25 per cent) but thereafter increases steadily. At Grades 10 and 11 the U. N. flag is actually higher than the U. S. flag. Thus there is a steady growth in appreciation of the U. N. flag to the point where it is about the same as that of the U. S.

5. *Animal Flags*

There were two flags with animals—Siam, showing elephants, and Iran, showing a lion. Whereas, Horowitz reported the Siamese flag highly chosen in Grade 1, there was no evidence in this investigation that perception of the animals led to a choice of these two flags.

D. Discussion

One question that might arise concerns the validity of choice of flags as a measure of attitude. It is more of a projective measure. Horowitz believed that flag choices would be correlated with attitudes. Our studies with high school and college students tend to confirm this. We doubt whether the average person can be objective when rating the beauty of flags even when the labels are missing. Flags are important symbols of patriotic feeling and often evoke emotional responses even when the immediate purpose of the symbol is no longer served. Thus we note that while most Americans do not associate the Red Cross organization with religion, the counterpart organization in Moslem countries is known as the Red Crescent, and in Israel, the Red Star of David.

In order to strengthen our contention that what is really being rated is indicative of attitudes toward the countries represented, we cite the findings from an earlier study (3) in which a panel of individuals with professional art training ranked the top flags as Greece, Saudi Arabia, Russia, Britain, and Norway. Brazil was rejected. Our *S*s certainly had no ranking approximating this. Because of the results with the art panel and the confirmatory evidence of significant correlations with a standard scale of nationalism (most of the correlations were significant) we feel that the flag scale is a reasonably valid instrument indicating feeling toward the object concerned.

The children in our sample apparently develop their attitudes at a substantially earlier age than those in the 1936 study. This is shown by the early acceptance of the U. S. flag, rejection of the Soviet. That the 1961 group is more sophisticated is shown in failure to generalize positive responses to the Liberian flag and failure to choose flags with animals on them. Perhaps our most interesting result is the steady growth in appreciation of the U. N. flag as the child goes through school until the U. N. flag is in the same range as that of the U. S. indicating a strong positive feeling toward the U. N. But why the crossover effect in Grades 10 and 11? Examination of the social studies offerings of both high schools shows that World History is taught in Grade 10, U. S. History in Grade 11. No unit of history is required in Grade 12. We can only speculate that the students in Grade 10 studying World History developed a positive attitude toward the United Nations which carried over to Grade 11 when U. S. History was being studied. In Grade 12, where history was not required and the instructor's influence was no longer present, the attitudes returned to the previous level.

How can we account for the differences between the two studies? There are several possible factors to be considered.

First, the 1961 sample was in an urban-suburban area in New York, the 1936 in rural Tennessee.

Second, the 1961 sample had a larger number of cases (1040 vs. 82) and covered a greater grade range (K-12 vs. 1-10). It was also based upon two school systems.

Third, the general cultural atmosphere in the country has brought about greater sensitivity to symbols such as flags. Since World War II there has been greater awareness of the position of world leadership and world influence held by the United States. The presence of the United Nations headquarters in New York City is clear evidence that the U. S. is not as remote from world affairs as it would have appeared to a rural Tennessee schoolboy in the mid-thirties. This greater awareness is reflected in the schools, in most homes, and in our organs of communication: newspapers, magazines, television.

On the basis of the three general points raised above we can explain some of the change in attitude. But there is a major question as to why the children in the 1961 sample showed their patterns of acceptance at such an early level, i.e., the rejection of the Soviet flag even in kindergarten. We feel that the best explanation of these early attitudes is in the influence of television. Today, according to one study (6) children watch television from a median of 20 hours a week in kindergarten to a median of 19 hours in the sixth grade. Undoubtedly, much of what the very young child views is over his head but he still takes in a great many values—from Captain Kangaroo featuring Skitch Henderson in a demonstration of classical instruments to Popeye and his inevitable spinach. Some of the children's programs feature a daily pledge of allegiance to the flag. News and other programs give the child of today far greater awareness of the world and of the United States than his father had twenty-five years ago.

E. Summary

Using an adaptation of the Horowitz test of patriotism, 1040 children in Grades K-12 were interviewed in two upstate New York communities. Compared with children interviewed twenty-five years earlier by Horowitz, children recognize flags at an earlier level; the United States flag is liked consistently at a high level, the Soviet rejected. This is apparent from even the earliest grades. Appreciation of the United Nations flag begins at a low level in kindergarten but shows a steady increase. At the high school level it is in the same range as the United States flag, although that of the United States does tend to be slightly higher. Several interpretations were offered to explain the data involving differences in location, size, and year of the two samples. An additional point was made concerning the role of television in helping to shape children's attitudes.

References

1. Ferguson, L. W. A revision of the primary social attitude scale. *J. of Psychol.,* 1944, **17**, 229-241.
2. Horowitz, E. L. Some aspects of the development of patriotism in children. *Sociometry,* 1941, **3**, 329-341.
3. Lawson, E. D. Attitude shift as related to palmar sweating in group discussion. Unpublished doctoral dissertation, Univ. Illinois, 1954.
4. Lawson, E. D., & Stagner, R. Group pressure, attitude change and autonomic involvement. *J. Soc. Psychol.,* 1957, **45**, 299-312.
5. Newcomb, T. M. Personality and Social Change. New York: Dryden, 1943.
6. Schramm, W., Lyle, J., & Parker, E. B. Television in the Lives of Our Children. Stanford, Calif.: Stanford Univ. Press, 1961.

Department of Psychology
State University of New York
Albany 3, New York

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 287-297.

TEMPORAL JUDGMENT AS A FUNCTION OF NOSOLOGICAL CLASSIFICATION AND EXPERIMENTAL METHOD*

University of Alabama, and Veterans Administration Hospital, Tuscaloosa

Joel S. Warm, James R. Morris, and John K. Kew[1]

A. Introduction

Although distortions of temporal awareness have been described in various neuropsychiatric disorders (e.g., 15, 20, 22), much of this material is based primarily upon clinical observation (25). Further, while some empirical investigations indicate differences between schizophrenics and normals in judgments of brief stimulus durations (8, 11, 16), except for a study by Dobson (5), little comparative data are available for other diagnostic groups.

Three methods are typically used to investigate the perception of brief stimulus intervals, *viz.*, production, reproduction, and verbal-estimation (1). Previous investigations using normal *O*s, point to the equivalence of these methods upon direct comparison (6, 9). A study by Clausen (2) indicates, however, that these methods may yield different results when judgments are made by schizophrenic *O*s. In his study, verbal-estimation judgments tended to be higher than those of production or reproduction while reproduction judgments appeared less reliable than those of the other methods.

Differences in stimulus presentation, mode of response available to the *O*, mental set required, and the relationship between personal and physical time (2) exist among the production, reproduction, and verbal-estimation techniques. Such differences may be related to the effects of psychopathology upon temporal experience. The present experiment attempted to determine if characteristic changes could be identified in the perception of brief stimulus durations as a function of nosological classification and these experimental methods.

B. Method

1. *Subjects*

Three groups of six male veterans each were selected: A normal control of hospital nursing assistants, a group of patients with the diagnosis of chronic

---

* Received in the Editorial Office on October 25, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors are grateful for the advice of Professors Earl C. Brown, Oliver L. Lacey, George E. Passey and Paul S. Siegel of the University of Alabama. Additionally, we are indebted to Drs. William B. Robinson and Charles W. Taylor, Tuscaloosa VA Hospital, for assistance in selecting the patient population.

undifferentiated schizophrenia, and another with the diagnosis of anxiety neurosis. Selection of patients was dependent upon psychiatric and psychological agreement regarding nosology. The patients selected were all privileged and none underwent any form of insulin or EST therapy during or immediately prior to experimentation. Three neurotic and two schizophrenic patients were receiving no medication during the experiment, the remainder were on drugs having a "depressant" action. A recent study (10) indicates that depressant drugs such as Chlorpromazine may have no significant effect on the perception of time when administered in clinical dosages under experimental conditions similar to those employed herein. All of the control *O*s were rated as being in relatively good health and none had any record of psychiatric illness.

There were no significant differences between groups on the variables of age, education, and WAIS vocabulary score. Each *O* was rated on a scale of cooperativeness adopted from Moran (18). Ratings were accomplished independently by two of the authors at each experimental session and averaged to yield a composite score. A Kruskal-Wallis one-way analysis of variance revealed no significant differences among the groups on this measure.

2. *Procedure*

Experimentation was carried out in a laboratory divided into two adjacent rooms set up into experimental and control sections. A Talk-A-Phone Model T-1M-5 two channel intercom system provided for one or two-way auditory communication at the discretion of *E*. Visual contact was provided by an observation window set in the wall partitioning the laboratory.

Temporal stimuli consisted of the duration of a 1,000 cps. tone presented binaurally via earphones at an intensity of 30db (*SL*). The output of a Beltone Model 12-A Audiometer served as the stimulus source. A rotary switch in series with the audiometer permitted *E* to control presentation of the tone by means of a Hunter Model 111-C Timer or *O* to turn the stimulus on via a response switch. A Standard Electric Model S-1 Precision Timer in parallel with the response switch recorded production and reproduction judgments.

*O*s participated in three experimental sessions making 20 judgments of each of four stimulus durations (0.5, 1, 5 and 10 seconds) under a different method in each session. Methods were counterbalanced and stimuli were randomly presented.

Verbal estimates were given orally. Production and reproduction responses were made by turning on the response switch. *O*s were instructed to attend to the stimulus durations, not switching behavior, and were told not to count or tap while the stimulus was on. *O*s were not informed as to the accuracy of their judgments.

C. RESULTS

Mean judgments as a function of method are presented for the patient groups and the normals in Figures 1-3. Generally, verbal-estimation judgments tend to be greater than, and production judgments less than, the physical standard. Reproduction tends to result in judgments above the physical standard at short intervals and judgments below this standard at longer intervals. These trends appear relatively consistent for each group.

Differences in response magnitude (judged time) for the various conditions were evaluated by analyses of variance. Results shown in Table 1 indicate significant differences between methods and intervals but no significant main effect. Further, there is a significant interaction between method and interval and a significant double order interaction.

TABLE 1
ANALYSIS OF VARIANCE OF RESPONSE MAGNITUDE AND VARIABILITY

Source of variation	Error term	df	Response magnitude Mean square	Response magnitude F	Response variability Mean square	Response variability F
1. Methods (M)	9	2	374.3460	29.88**	72.5589	14.05**
2. Nosology (N)	8	2	53.4039	1.90	12.5162	1.44
3. Interval (I)	9	3	1258.1640	100.43**	65.0185	12.59**
4. M × N	8	4	68.1919	2.43	12.7034	1.46
5. M × I	9	6	126.1122	10.07**	28.9462	5.61**
6. N × I	9	6	19.8401	1.58	6.1993	1.20
7. M × N × I	9	12	25.2951	2.02*	6.1388	1.19
Total between cells		35				
8. Between subjects in same group		45	28.0770		8.6840	
9. Interaction: pooled subjects × trials		135	12.5282		5.1639	
Total within cells		180				
10. Total		215				

* $p < .05$.
** $p < .01$.

Differences among methods were evaluated by *t* test for correlated samples. Results indicate that judged time under verbal-estimation is significantly higher than under production ($t = 7.41$, $p < .01$) or reproduction ($t = 5.98$, $p < .01$), but that the latter two do not differ significantly from each other.

Inasmuch as no significant effect was found for nosology, data for all *O*s were grouped and indifference intervals graphically determined for production and reproduction. These values are .20 seconds and 3.48 seconds respectively.

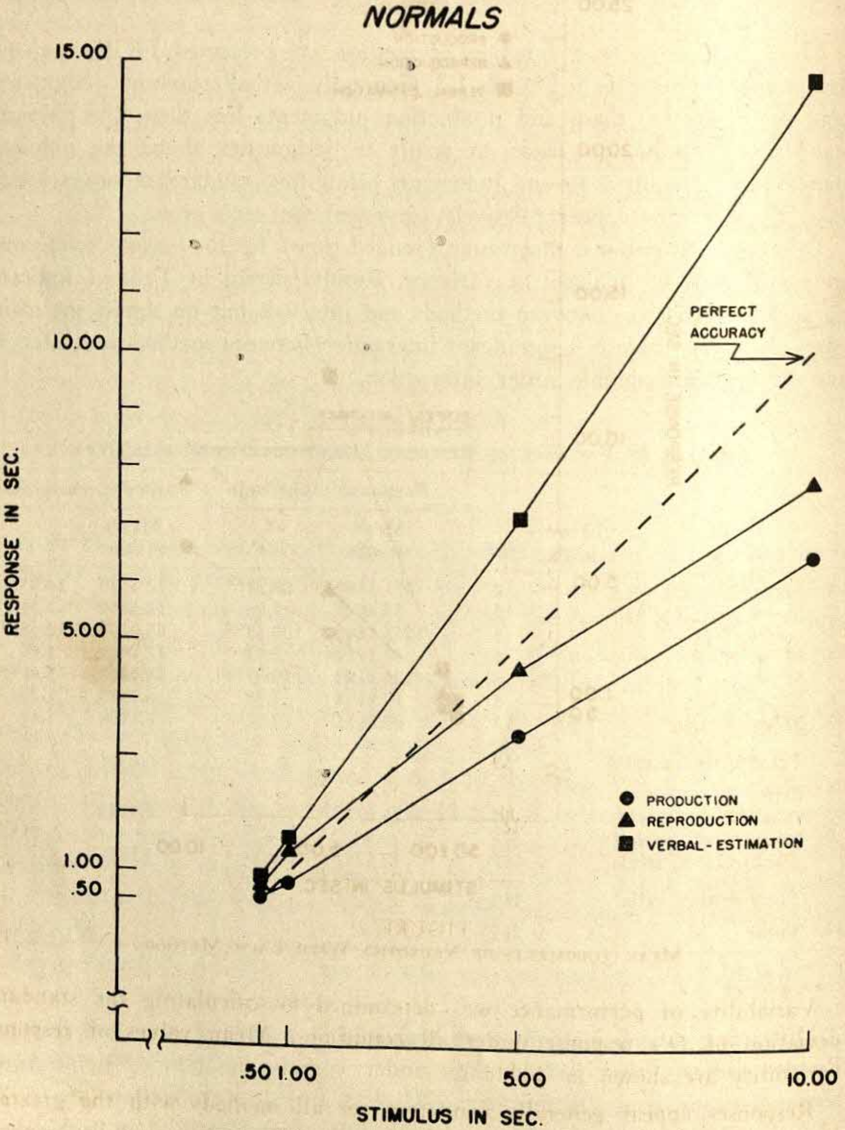


FIGURE 1
MEAN JUDGMENTS OF NORMALS WITH EACH METHOD

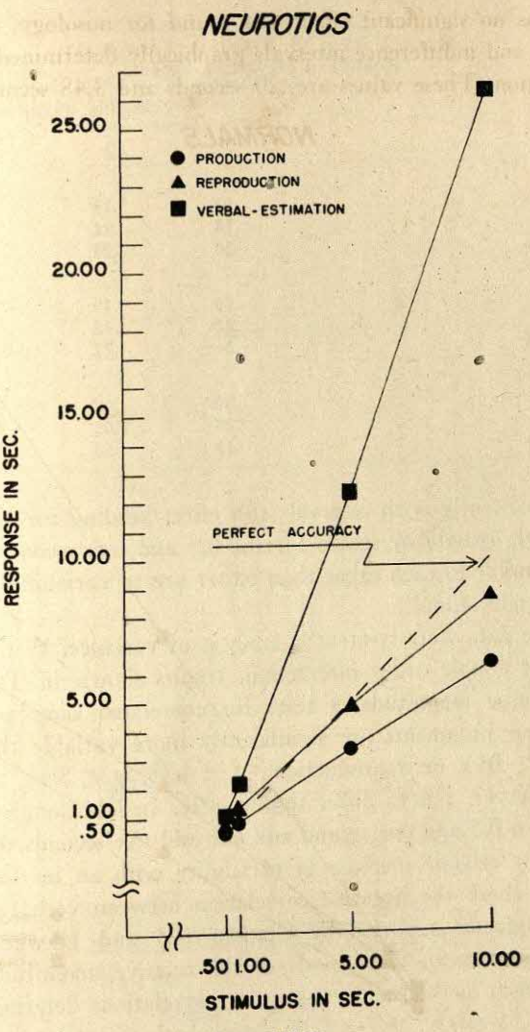


FIGURE 2
MEAN JUDGMENTS OF NEUROTICS WITH EACH METHOD

Variability of performance was determined by calculating the standard deviation of *O*'s response under all conditions. Mean values of response variability are shown in Table 2.

Responses appear generally consistent for all methods with the greatest variation being associated with verbal-estimation. Response variability appears

TABLE 2
MEAN VALUES OF RESPONSE VARIABILITY
(Standard Deviation Scores)

	Stimulus durations (Sec.)			
	.50	1	5	10
Production				
Normals	.10	.18	.47	.74
Neurotics	.13	.21	.55	.84
Schizophrenics	.20	.28	.76	.89
Reproduction				
Normals	.16	.19	.64	.82
Neurotics	.22	.32	.78	1.17
Schizophrenics	.24	.27	.84	1.39
Verbal-estimation				
Normals	.20	.42	2.39	3.92
Neurotics	.20	1.08	3.22	10.20
Schizophrenics	.13	.56	1.33	3.37

to increase consistently with interval; this effect holding for all methods and groups. Further, variability scores for the 0.5 and one second stimuli appear more nearly similar to each other than either are to variability scores for the five and 10 second stimuli.

The response data were treated by analysis of variance. With the exception of a significant double order interaction, results shown in Table 1 parallel those for response magnitude. t tests for correlated samples indicate that verbal-estimation judgments are significantly more variable than production ($t = 4.75$, $p < .01$) or reproduction ($t = 4.38$, $p < .01$) responses. The latter two, however, fail to differ significantly. In addition, except for comparisons between 0.5 and one second and one and five seconds there appears to be a statistically reliable increase in variability with an increase in interval.

In order to check the negative correlation between verbal-estimation and production judgments reported by Clausen (2) and Hawkes, Bailey, and Warm (9), comparisons were made of the relative magnitude of each *O*'s response with each method and rank order correlations determined. A significant rho of —.51 ($df = 16$, $p < .05$) was obtained between production and verbal-estimation. Comparisons between production vs. reproduction and reproduction vs. verbal-estimation yielded rhos of .35 and —.21 respectively. These values are not statistically reliable.

In view of Kleiser's (14) report that reproduction of temporal intervals become progressively longer over trials as *O*s manifest improvement in the experimental task, comparisons were made of *O*'s first and last ten judgments under each method at each interval. Data for all *O*s were grouped. Results

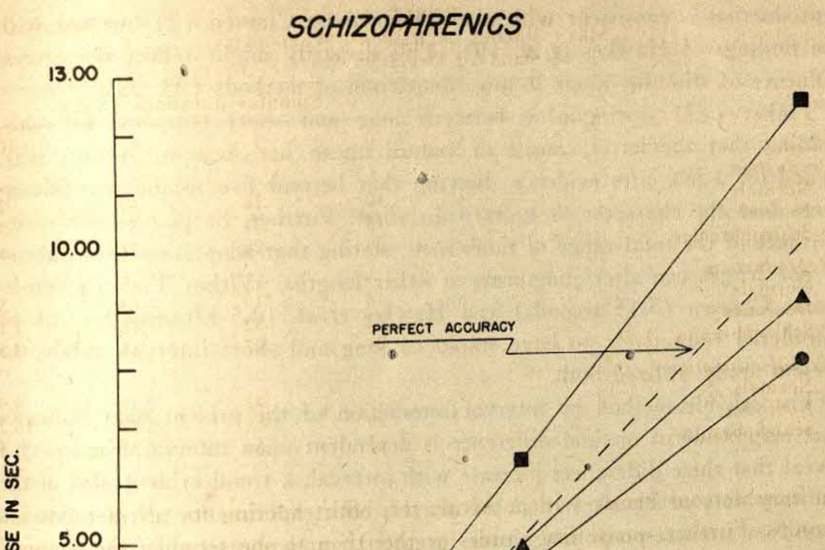


FIGURE 3
MEAN JUDGMENTS OF SCHIZOPHRENICS WITH EACH METHOD

show a statistically reliable increase in magnitude of reproduction judgments during the last ten trials at all intervals. No such increase was evident for production or verbal-estimation judgments.

D. DISCUSSION

The finding of significantly higher response magnitudes resulting from verbal-estimation judgments when compared with those of production and reproduction is consistent with a trend noted by Clausen (2) but not with the findings of Hawkes *et al.* (9). This disparity might reflect the crucial influence of stimulus range in any comparison of methods (23, 28).

Tinker (23) distinguishes between long and short temporal intervals, holding that the latter consist of stimuli up to five seconds. Additionally, Woodrow (28) cites evidence showing that beyond five seconds continuous tones lose the character of unitary duration. Further, he points to the importance of the total range of times used, stating that adaptation to durations of one length can alter judgments at other lengths. Within Tinker's framework, Clausen (5-15 seconds) and Hawkes *et al.* (0.5-4.0 seconds) can be considered respectively to have employed long and short intervals while the present study utilized both.

The reliable method by interval interaction of the present data indicates that magnitude of method difference is dependent upon interval. Figures 1-3 reveal that these differences increase with interval, a trend evident also in the auditory data of Hawkes *et al.* Moreover, both experiments reveal points of transition from response magnitudes greater-than to shorter-than the physical standard (indifference interval) for production and reproduction but not for verbal-estimation. This suggests that extension of the data of Hawkes *et al.* to longer intervals would lead to the expectancy of methodological differentiation.

Attribution of such differences to the effects of increased interval gains further credence from Woodrow's (28) statement that the indifference interval tends in part to be a function of characteristics of the temporal stimuli presented. The values of .20 and 3.48 seconds for production and reproduction used here differ from those of Hawkes *et al.* whose values fall between one and two seconds. The values of the indifference interval in this study, however, as well as those of Hawkes *et al.*, fall within the typical range quoted by Woodrow, *viz.,* from under .36 to 5.0 seconds. Additionally, overestimation of short intervals and underestimation of long intervals with production and reproduction agrees with a general trend described by Wallace and Rabin (25).

A significant difference in response magnitude between methods is also inconsistent with the findings of Gilliland and Humphries (6). However, these investigators utilized open interval stimuli while closed intervals were employed here. The equivalence of these types of temporal stimuli remains questionable (25).

The significant negative rank order correlation between production and verbal-estimation judgments and the lack of correlation with reproduction lends strong support to Clausen's hypothesis that the relation between subjective and objective time is measured inversely by production and verbal-estimation while a different process may underlie reproduction performance.

An increase in reproduction response magnitude over trials agrees with the results of Kleiser (14) and DeWolfe and Duncan (4). Temporal estimates have been found to vary directly with boredom (17). The present finding may reflect an increase in lassitude during the repetitive performance of a non-rewarding task and consequent overestimation of the standard. Failure of judgments to increase with trials under production and verbal-estimation is a further indication that different processes may be involved in these methods as compared with reproduction.

Considering response variability, the general consistency of performance seen here agrees with a number of other investigations (7, 9, 21). Additionally, a trend toward increase in variability with an increase in stimulus duration agrees with the findings of Dobson (5) and lends further support to the crucial role of stimulus range in time estimation. Significantly greater variability of verbal-estimation judgments compared with production and reproduction is inconsistent, however, with the results of both Hawkes *et al.* (9) and Clausen (2). In view of the trend toward increased variability at longer intervals and a significant method by interval interaction, the discrepancy between the present data and that of Hawkes *et al.*, may be attributed plausibly to differences in stimulus range. There appears no ready explanation for the lack of agreement between the present data and that of Clausen. It should be noted, however, that Clausen's reliability coefficients were computed from comparisons of visual and auditory judgments. It is possible that the reliability of heteromodal judgments is poorer for reproduction than for the other methods.

Failure to obtain significant differences as a function of nosology confirms the findings of Cutler (3) and Dobson (5) but disagrees with a number of other studies (8, 11, 16, 19). Taken as a group, these last are characterized by methodological heterogeneity. Recently, Winder (27) pointed out that deficits in the schizophrene tend to be evidenced primarily in social situations. Moreover, a study by Vandenburg (24) demonstrated the difficulty of differentiating between neurotics and schizophrenics on a number of perceptual tasks, and Webster, Goldstone and Webb (26) have indicated that disparities in the temporal judgments of normals and schizophrenics may in part be due to differences in psychophysical methodology. Considering these findings, the present outcome is not surprising. Perhaps equivocality regarding the effects of psychopathology on temporal experience reflects the sensitivity of temporal perception to variations in experimental procedure as well as lack of precision in the present diagnostic scheme (12, 13).

E. Summary

An attempt was made to identify characteristic changes in the perception of brief stimulus intervals as a function of nosological classification and experimental method. The techniques of production, reproduction, and verbal-estimation were used with stimulus durations of 0.5 to 10.0 seconds. Judgments were made by groups of normal, neurotic, and schizophrenic *O*s.

Significant differences were found for method but not for nosology. Verbal estimation judgments tended to be higher and more variable than those of the other methods. Further, the magnitude of production and verbal-estimation judgments tended to be negatively correlated and unrelated to reproduction judgments. This was taken as confirming a hypothesis of Clausen, *viz.*, that production and verbal-estimation measure inversely the relation between subjective and objective time, while reproduction may be dependent upon other processes.

References

1. Bindra, D., & Waksberg, H. Methods and terminology in studies of time estimation. *Psychol. Bull.*, 1956, **53**, 155-159.
2. Clausen, J. An evaluation of experimental methods of time judgment. *J. Exp. Psychol.*, 1950, **40**, 756-761.
3. Cutler, H. M. The effects of stress on the estimation of short time intervals by neurotic and normal subjects. Unpublished doctoral dissertation, Purdue Univ., 1951.
4. DeWolfe, R. K. S., & Duncan, C. P. Time estimation as a function of level of behavior of successive tasks. *J. Exp. Psychol.*, 1959, **58**, 153-158.
5. Dobson, W. R. An investigation of various factors involved in time perception as manifested by different nosological groups. Unpublished doctoral dissertation, Purdue Univ., 1951.
6. Gilliland, A. R., & Humphries, D. W. Age, sex, and interval as variables in time estimation. *J. Genet. Psychol.*, 1943, **63**, 123-130.
7. Gridley, P. F. The discrimination of short intervals of time by fingertip and by ear. *Amer. J. Psychol.*, 1932, **44**, 18-43.
8. Guertin, W. H., & Rabin, A. I. Misperception of time in schizophrenia. *Psychol. Rep.*, 1960, **7**, 57-58.
9. Hawkes, G. R., Bailey, R. W., & Warm, J. S. Method and modality in judgments of brief stimulus duration. *J. Aud. Res.*, 1961, **1**, 133-144.
10. Hawkes, G. R., Joy, R. J. T., & Evans, W. O. Autonomic effects on estimates of time: Evidence for a physiological correlate of temporal experience. *J. of Psychol.*, 1962, **53**, 183-191.
11. Johnston, H. M. A comparison of time estimation of schizophrenic patients with that of normals. Unpublished Master's thesis, Brown Univ., 1939.
12. King, G. F. Research with neuropsychiatric samples. *J. of Psychol.*, 1954, **38**, 383-387.

13. KLEIN, D. B. Abnormal Psychology. New York: Holt, 1951.

14. KLEISER, J. R. The effect of habit formation, mental activity, and gross muscle movements on the reproduction of time. *Dissertation Abstr.*, 1953, **13**, 886.

15. LEWIS, A. The experience of time in mental disorder. *Proc. Roy. Soc. Med.*, 1931, **25**, 611-620.

16. LHAMON, W. T., & GOLDSTONE, S. The time sense. A.M.A. *Arch. Neurol. Psychiat.*, 1956, **76**, 625-629.

17. LOEHLIN, J. C. The influence of different activities on the apparent length of time. *Psychol. Monogr.*, 1959, **73**, No. 4 (Whole No. 474).

18. MORAN, T. J. Vocabulary knowledge and usage among normal and schizophrenic subjects. *Psychol. Monog.*, 1953, **67**, No. 20 (Whole No. 370).

19. RABIN, A. I. Time estimation of schizophrenics and non-psychotics. *J. Clin. Psychol.*, 1957, **13**, 88-90.

20. SCHILDER, P. Psychopathology of time. *J. Nerv. Ment. Dis.*, 1936, **83**, 530-545.

21. SIEGMAN, A. W. Anxiety, impulse control, intelligence, and the estimation of time. *J. Clin. Psychol.*, 1962, **28**, 103-105.

22. STRAUSS, E. C. Disorders of personal time in depressive states. *South. Med. J.*, 1947, **40**, 254-259.

23. TINKER, M. A. Temporal perception. *In* Boring, E. G., Langfeld, H. S., & Weld, H. P. (*Eds.*), *Psychology*. New York: Wiley, 1935. Pp. 246-259.

24. VANDENBERG, S. G. Differences between neurotics, psychotics and normals on perceptual tests: Reanalysis of Eysenck's results. *J. Clin. Psychol.*, 1959, **15**, 373-376.

25. WALLACE, M., & RABIN, A. I. Temporal experience. *Psychol. Bull.*, 1960, **57**, 214-236.

26. WEBSTER, F., GOLDSTONE, S., & WEBB, W. W. Time judgment and schizophrenia: Psychophysical method as a relevant contextual factor. *J. of Psychol.*, 1962, **54**, 159-164.

27. WINDER, C. L. Some psychological studies of schizophrenics. *In* Jackson, D. D. (*Ed.*), *The Etiology of Schizophrenia*. New York: Basic Books, 1960. Pp. 191-247.

28. WOODROW, H. Time perception. *In* Stevens, S. S. (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951. Pp. 1224-1236.

VA Hospital
Tuscaloosa, Alabama

Published as a separate and in The Journal of Psychology, [illegible]

THE RELATIONSHIP BETWEEN PERSISTENCE AND PERSONALITY[*]

[illegible]

H. M. [illegible] AND C. G. [illegible]

A. INTRODUCTION

Relatively little work has been done on the relationship between persistence and personality. The studies that have been done have [illegible] physical persistence [illegible] factor analysis of a large number of persistence tests that there was [illegible] of persistence — mental and physical persistence — suggests the need to use both mental and physical persistence tests in studying the relationships between persistence and personality or any other variable.

With regard to the first [illegible] Eysenck [illegible] predicted on the basis of [illegible] theory [illegible] that [illegible] negative [illegible] for [illegible] extravert's physical [illegible] greater than the [illegible] obtained [illegible] Cattell and Eysenck [illegible] persistent [illegible] prediction was [illegible] has found however that [illegible] persistent than [illegible]

The studies of the relationship between sex and physical persistence as measured by the hand dynamometer have found that females are more persistent than males [illegible] though no difference was found between these groups in the study by Cattell and Eysenck. No significant differences have been found between normals and neurotics when the leg persistence test is used [illegible]

There is no evidence with regard to the relationship between [illegible] and mental persistence. Insofar as physical persistence tests can be regarded as simple motor tasks and mental [illegible] can be regarded as [illegible]

---

[*] Received in the Editorial Office on [illegible] and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[illegible] work reported in this paper was supported in part by a grant from Wallace Laboratories [illegible]

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 299-306.

THE RELATIONSHIP BETWEEN PERSISTENCE AND PERSONALITY*[1]

Regina General Hospital, Saskatchewan, Canada

H. M. BRACHMAN AND C. G. COSTELLO

A. INTRODUCTION

Relatively little work has been done on the relationships between persistence and personality. The studies that have been done have used mainly physical persistence tasks (6). MacArthur's finding (11) in his factor analysis of a large number of persistence tasks that there were two factors of persistence —mental and physical persistence—suggests the need to use both mental and physical persistence tasks in studying the relationships between persistence and personality or any other variables.

With regard to physical persistence, Costello and Eysenck (2) predicted on the basis of Eysenck's general theory (4, 5) that pain would be a stronger negative drive for introverts than for extraverts thus resulting in the extravert's physical persistence being greater than that of introverts. Data obtained on a hand dynamometer, in line with this prediction, were presented.

Costello and Eysenck also predicted that boredom would be a stronger negative drive in extraverts who would consequently be expected to be less persistent than introverts on mental tasks. No data in connection with this prediction was presented. Foulds (9) has found however, that introverts were more persistent than extraverts on Raven's Progressive Matrices.

Two studies of the relationship between neuroticism and physical persistence as measured by the hand dynamometer have found that normals are more persistent than neurotics (7, 10) though no difference was found between these groups in the study by Costello and Eysenck. No significant differences have been found between normals and neurotics when the leg persistence test is used (3, 7, 10).

There is no evidence with regard to the relationship between neuroticism and mental persistence. Insofar as physical persistence tasks can be regarded as simple motor tasks and mental persistence tasks can be regarded as rela-

---

* Received in the Editorial Office on October 26, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The work reported in this paper was supported in part by a grant from Wallace Laboratories, New Jersey.

tively complex ideational tasks, one would predict on the basis of the Yerkes-Dodson Law (12) that differences between normals and neurotics in favour of normals would be greater on mental persistence tasks than on physical persistence tasks. The argument underlying this prediction is that the high drive of the neurotics would interfere to a greater extent with performance on the complex task.

The relationship between persistence and other experimental variables has also been neglected though they would appear to be promising areas of study. There is no evidence on the effects of individual vs. group conditions on persistence though predictions have been made by Costello and Eysenck (2). It has been predicted that extraverts will increase in persistence to a greater extent than introverts when group conditions of testing are compared with individual conditions of testing. This prediction is also based on Eysenck's general theory, the argument being that since extraverts condition less readily than introverts, they would be expected to form the conditioned responses underlying socialization less readily than introverts. Thus the extraverts would be more subject to the changing motivational conditions of the moment and would be more dependent on the judgments and values of the immediate group.

Costello and Eysenck (2) did not find a significant difference between males and females in physical persistence. There is no evidence relating to sex differences in mental persistence.

The experiment to be reported was designed to investigate further the relationships between physical and mental persistence and the variables discussed above. Specific hypotheses were: (*a*) the extraverts' physical persistence will be greater than that of introverts; (*b*) introverts will be more persistent on a mental task than extraverts; (*c*) extraverts will increase in persistence on mental and physical tasks when group conditions are compared with individual testing conditions to a greater extent than introverts; (*d*) normals will be more persistent than neurotics on a mental task.

Because of the conflicting data, no prediction was made with regard to the effects of neuroticism on physical persistence, or in relation to sex differences, or on the interaction between sex neuroticism and individual vs. group conditions. All these variables and their interactions were, however, investigated.

B. PROCEDURE

1. *Subjects*

*S*s were 80 school children between the ages of 14 years and 16 years selected from a group of 538 children in the same age range on the basis of their extraversion and neuroticism scores on the Junior M.P.I. (8). The Mean E score of the 538 *S*s was 12.05 SD 3.61 and the Mean N score was 5.90 SD 2.84. The 80 *S*s were split into eight groups of 10 *S*s thus: Group 1—Male extraverted normals with mean E score of 15.5 and mean N score of 2.4; Group 2—Female extraverted normals, E = 15.0 N = 2.9; Group 3—Male extraverted neurotics, E = 15.8 N = 9.2; Group 4—Female extraverted neurotics, E = 16.1 N = 10.2; Group 5—Male introverted normals, E = 7.6 N = 2.9; Group 6—Female introverted normals, E = 7.7 N = 3.2; Group 7—Male introverted neurotics, E = 7.1 N = 9.5; Group 8—Female introverted neurotics, E = 7.2 N = 10.3.

Each *S* in the eight groups of 10 *S*s was allotted randomly to one of two groups. *S*s in Group 1 were tested under individual conditions of testing and those in Group 2 under group conditions.

2. *Tests*

The hand dynamometer was used as a test of physical persistence. The length of time *S* could maintain the pointer just beyond a point which was 2/3 of his strength was taken as his physical persistence score. This test was found by MacArthur (11) to have a factor loading of 0.49 on a factor defined as physical persistence.

The mental persistence test consisted of an unsolvable Magic Number Square. *S*'s mental persistence was the length of time he continued to work on the Magic Number Square when he was regularly given the opportunity to switch to a Magic Word Square. This test was found by MacArthur (11) to have a loading of 0.62 on a factor defined as mental persistence.

3. *Method*

The order of presentation of the tests was the same for all *S*s—physical persistence followed by mental persistence tests.

All *S*s tested individually were given 6 trials of strength on the hand dynamometer with 20-second intervals between trials using his right and left hands on alternate trials. Two-thirds of the mean strength of each hand was calculated and after a one-minute rest interval *S* was given two trials of persistence at this setting, *S* being instructed to keep the pointer as long as possible at this setting for his right hand and at this setting for his left hand with his left hand. The two trials correlated 0.83 and their mean was *S*'s physical persistence score.

S was then given the test of mental persistence. He was instructed to put numbers in the nine cells of a 3 × 3 square using only numbers from one to nine, not repeating any numbers and placing them so that the rows and columns all added up to the same amount. After five minutes *S* was given the opportunity to switch to a Magic Word Square and thereafter was given the same invitation at 10-minute intervals. The length of time before *S* switched to the Magic Word Square was his mental persistence score.

The procedure was exactly the same for *S*s in the group conditions. The order in which *S*s were given the physical persistence tests was determined randomly and *S* performed on the tests whilst being watched by the other members of his group. All *S*s in each group of five *S*s were given the mental persistence task at the same time.

C. RESULTS

The results of an analysis of variance performed on the data from the physical persistence task are presented in Table 1.

TABLE 1
ANALYSIS OF VARIANCE ON DATA FOR PHYSICAL PERSISTENCE

SOURCE	*df*	Sum Squares	M.S.V.	*F*	*p*
Total	79	5071.49			
Sex	1	1419.61	1419.61	32.85	$< .001$
Ext.-Int.	1	25.31	25.31	—	NS
Neuroticism	1	.11	.11	—	NS
Ind.-Gp	1	270.11	270.11	6.25	$< .025$
Sex/E	1	5.52	5.52	—	NS
Sex/N	1	2.82	2.82	—	NS
Sex/Gp	1	285.02	285.02	6.60	$< .025$
E/N	1	37.82	37.82	—	NS
E/Gp	1	.12	.12	—	NS
N/Gp	1	43.52	43.52	—	NS
Sex/E/N	1	.60	.60	—	NS
Sex/E/Gp	1	108.10	108.10	2.52	NS
E/N/Gp	1	40.60	40.60	—	NS
Residual	69	2981.53	43.21		

It will be seen that three significant *F* ratios emerge, namely those for sex, individual vs. group conditions and the sex/individual vs. group interaction. The results of *t* tests based on the analysis of variance indicate that: (*a*) males are significantly more persistent than females, the means being males—42.15 seconds; females—25.30 seconds; (*b*) group conditions of testing resulted in an overall increase in physical persistence when compared with individual conditions of testing, the mean persistence scores being group conditions: 37.40 seconds; individual conditions: 30.05 seconds; (*c*) there is a significant increase in persistence when the mean score of males under group conditions is compared with that of males under individual conditions.

The increase in the scores of the females from group to individual conditions is not significant. The means are: Males—Group Conditions, $M = 99.20$ seconds; Individual Conditions, $M = 69.40$ seconds; Females—Group Conditions, $M = 50.40$ seconds; Individual Conditions, $M = 50.80$ seconds.

The results of the analysis of variance on the data from the mental persistence task are presented in Table 2.

TABLE 2
ANALYSIS OF VARIANCE ON DATA FOR MENTAL PERSISTENCE

SOURCE	df	Sum squares	M.S.V.	F	p
Total	79	88,499,091.69			
Sex	1	7,085,856.01	7,085,856.01	17.52	$< .001$
Ext.-Int.	1	86,922.11	86,922.11	—	NS
Neuroticism	1	8,539,284.61	8,539,284.61	21.12	$< .001$
Ind.-Gp.	1	28,394,636.51	28,394,636.51	61.02	$< .001$
Sex/E	1	428,805.62	428,805.62	—	NS
Sex/N	1	1,237,282.52	1,237,282.52	3.06	NS
Sex/Gp.	1	4,302,672.62	4,302,672.62	10.64	$< .005$
E/N	1	542,686.52	542,686.52	—	NS
E/Gp.	1	248,979.62	248,979.62	—	NS
N/Gp.	1	6,784,377.62	6,784,377.62	16.78	$< .001$
Sex/E/N	1	2,949,504.00	2,949,504.00	7.47	$< .01$
Sex/E/Gp.	1	836,609.50	836,609.50	2.12	NS
E/N/Gp.	1	989,457.60	989,457.60	2.50	NS
Residual	69	27,898,083.93	404,320.06		

It will be seen that six significant F ratios emerge, namely those for sex, neuroticism, individual vs. group conditions, sex/individual vs. group interaction, neuroticism/individual vs. group interaction, sex/extraversion/neuroticism interaction.

The results of t tests based on the analysis of variance indicate that (*a*) females are more pesistent on the mental persistence task than males—female mean score $= 57.04$ minutes; male mean $= 37.20$ minutes; (*b*) normals are more persistent than neurotics—normal mean $= 58.01$ minutes; neurotic mean $= 36.23$ minutes; (*c*) group conditions resulted in an overall increase in persistence—group mean $= 66.98$ minutes; individual mean $= 27.26$ minutes; (*d*) the persistence of the females under group conditions is significantly greater than that of females under individual conditions. The difference in the case of the males is not significant.

The means are: Females—group conditions, $M = 42.31$ minutes; individual conditions, $M = 14.72$ minutes; Males—group conditions, $M = 24.66$ minutes; individual conditions, $M = 12.53$ minutes; (*e*) The difference between group conditions and individual conditions in favour of group conditions is significant in the case of normals. The difference in the case of neurotic

*S*s is not significant. The means are: Normals—group conditions, $M = 43.79$ minutes; Normals—individual conditions, $M = 14.22$ minutes; Neurotics—group conditions, $M = 23.19$ minutes; individual conditions, $M = 13.04$ minutes.

Although the strength factor in the physical persistence task is controlled to a large extent by taking 2/3 of each *S*'s strength, in view of the significant sex effect it was decided to correlate the strength and persistence scores for the total *N*. The product moment correlation was $r = —0.14$ which is insignificant.

D. DISCUSSION

The experiment did not produce evidence in line with the hypothesis that extraverts would be more persistent than introverts on a task of physical persistence and thus failed to confirm the findings of Costello and Eysenck (2). This suggests the possibility that the significantly greater persistence of the extraverts when compared with introverts in the English study may be due to factors other than physiological ones. If indeed the difference found in Costello and Eysenck's study was due to a factor or factors such as differences in level of motivation aroused by the task, then the failure to confirm the findings in the present study may be due to cultural differences. The increase in physical persistence produced by group conditions of testing certainly indicates that factors other than physiological are playing a role. Furthermore, there is some evidence based on item analyses of the Junior Maudsley Personality Inventory that English and Canadian children may differ in their expression of extraversion and introversion (1).

The data also failed to confirm the prediction that introverts would be more persistent on a mental persistence task and the extraverts would increase their persistence to a greater extent when group conditions of testing were compared with individual conditions of testing. Only further experiments will enable us to say at what stage from general theory to experimental testing there is a weakness.

The findings were in line with the hypothesis that normals would be more persistent than neurotics on a mental task. No differences between these two groups were found on the physical persistence task. We suggested in the introduction that the interference of neuroticism would be greater on the mental persistence task because of its greater complexity. It could of course be argued that if anxiety must build up to a certain level before it interferes with persistence then the difference in the average time spent on the mental task (23.55 minutes) and physical task (16.00 seconds) may be an important factor. Here again only further experimentation will give us the answer.

The greater increase in persistence produced by group conditions in the case of the normal *S*s lends itself to numerous possible explanations the listing of which without supporting data would be of little value.

Costello and Eysenck (2) did not find a significant sex difference in physical persistence. In the present study it was found that males were more persistent on the hand dynamometer than females. This suggests the possibility of a cultural difference in attitude to the persistence task. That the males in the present study were perhaps more ego involved in the task is suggested by the greater increase in their persistence from individual testing conditions to group conditions of testing.

No data on English children in relation to mental persistence is at hand with which to compare our present finding of the greater mental persistence of females than males and the greater increase in persistence of females from individual to group testing conditions.

E. Summary

It was found, as predicted, that normal *S*s were more persistent on a test of mental persistence than neurotics—neuroticism being defined by scores on the neuroticism scale of the Jnuior Maudsley Personality Inventory. It was also found that normals increased their mental persistence to a greater extent than neurotics when group conditions of testing were compared with individual testing conditions.

Males were found to be more persistent than females on a test of physical persistence and to increase in physical persistence to a greater extent when group conditions for testing were compared with individual testing conditions.

Females were more persistent than males on a test of mental persistence and increased their mental persistence to a greater extent than males when group conditions of testing were compared with individual testing conditions.

Contrary to prediction, the extraversion-introversion variable did not have an effect on physical or mental persistence either when acting alone or when interacting with the other variables studied.

References

1. Costello, C. G., & Brachman, H. M. Cultural and sex differences in extraversion and neuroticism reflected in responses to a Children's Personality Inventory. *Brit. J. Educ. Psychol.* (In press).
2. Costello, C. G., & Eysenck, H. J. Persistence, personality and motivation. *Percept. & Mot. Skills*, 1961, **12**, 169-170.

3. EYSENCK, H. J. The Scientific Study of Personality. London: Routledge & Kegan Paul, 1952.
4. ———. Dynamics of Anxiety and Hysteria. New York: Praeger, 1957.
5. EYSENCK, H. J. (*Ed.*) Experiments in Personality. (2 Vols.) London: Routledge & Kegan Paul, 1960.
6. ———. Handbook of Abnormal Psychology. New York: Basic Books, 1961.
7. EYSENCK, S. B. G. A dimensional analysis of mental abnormality. Unpublished Ph.D. Thesis, Univ. London Lib., 1955.
8. FURNEAUX, W. D., & GIBSON, H. B. A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 1961, **31**, 204-207.
9. FOULDS, G. A. The ratio of general intellectual ability to vocabulary among psychoneurotics. *Internat. J. Soc. Psychiat.*, 1956, **1**, 5-12.
10. HIMMELWEIT, H. T., & PETRIE, A. The measurement of personality in children, an experimental investigation of neuroticism. *Brit. J. Educ. Psychol.*, 1951, **21**, 9-29.
11. MACARTHUR, R. S. An experimental investigation of persistence in secondary school boys. *Canad. J. Psychol.*, 1955, **9**, 42-54.
12. YERKES, R. M., & DODSON, J. D. The relation of strength of stimulus to rapidity of habit formation. *J. Comp. Psychol.*, 1908, **18**, 459-482.

Regina Psychological Laboratory,
Munroe Wing, Regina General Hospital,
Regina, Saskatchewan.

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 307-311.

STUTTERING AND SELF CONCEPT*

Rutgers University, Newark, New Jersey

ROBERT W. RIEBER

A. INTRODUCTION

The literature relating to stuttering and self concept has been approached from two major directions. The first seeks to find experimentally if stutterers differ from normals in body image or self concept, and the second deals with the clinical phenomena of disturbed self concept as related to the symptomatology and treatment of stuttering. The purpose of this article is to discuss and clarify the issues involved in both of these areas.

B. THEORETICAL CONSIDERATIONS

One basic question that seems to be critical in this area is the nature of the relationship of stuttering to the stutterer's self concept. Van Riper (16) states, "If the self could be equated with stuttering then by working with the stuttering we might be able to alter the self." Shearer (14) apparently agrees with Van Riper but goes even further and states, "There must be conflicting body images present with the stutterer since no one stutters all the time or to every auditor." In terms of therapy, Shearer (14) again agrees with Van Riper when he states,

> During the process of therapy we speak of helping the stutterer to learn to "live with himself" and "face the stuttering." Thus even our language in referring to therapy denotes a program of synthesizing the dichotomy of conflicting self concepts and body images into a harmonious resolution of the stutterer's two selves.

There seems to be two theoretical assumptions explicitly expressed in Shearer's comments. The first is that stuttering and self concept are equal, and the second is that stutterers have two conflicting selves, one stuttering and the other not stuttering.

It is quite difficult to accept these two assumptions because of two important factors. The first point is that psychological theory and research dealing with the self suggest that self concept or body image be understood as a system of holistic unity and not as simply additive or elemental in nature.

---

* Received in the Editorial Office on October 30, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

The research of Rogers (13), Fisher and Cleveland (8), Witkin (18), and the theories of Lewin (12), Adler (1), Allport (2), and Dorsey (5), support this contention. The second point is that any behavior stuttering included should be understood as an extension or manifestation of the self and not as something alien to or dichotomized from the self.

With this frame of reference, stuttering is understood as an intrinsic part of the self or body image of the person with a stuttering problem. The stutterer has but one self which may be in varying degrees of conflict on different levels of abstraction. The theory and research of Sheehan (15) seems to support this point of view. An important point here is that this conflict is an intraself conflict and not a dichotomized self conflict. If this is the case, then stuttering therapy need not be concerned with the incorporation of stuttering into the body image of the stutterer; for assuredly, it was there to begin with. It may be internalized as some authorities believe or temporarily denied in some fashion, Douglas and Quarrington (6).

When a stutterer perceives himself without the tendency for stuttering as in those conditions described by Bloodstein (3), it may be said that the stutterer is experiencing a phenomenon very similar to what has been called self alienation, Fromm (10). This acting *as if* he is not himself may be understood as a temporary reduction or relief from the conflict which results in a certain degree of false fluency. Shearer (14) feels that the weakest point in our stuttering therapy is the relapse and attempts to describe it as follows:

> Perhaps the weakest point in our stuttering therapy today—namely, the relapse, can be more easily understood in terms of the self concept and body image. If the stutterer is able to see his stuttering tendency as a permanent part of his self concept and body image, he can be aware of the subtle cues in his environment which tend to set off and perpetuate the tremors in his blocks. If, on the other hand, he perceives himself as a person without the tendency for stuttering, he will not acknowledge his mild blocks or environment cues, and his stuttering can once again thrive on avoidance and repression.

The phenomenon of relapse as related to self concept seems to be better understood if we relate self-alienating behavior, or what Shearer calls avoidance and repression, to relapse as poles of a single dimension. It may be that the greater the self-alienating behavior the greater the chance for a consequent relapse into stuttering. Clinical experience seems to bear this out. Relapse and self-alienating behavior can be understood as a manifestation of the total self vacillating in a state of conflict. The factors that tend to set off or perpetuate stuttering can be seen as distortions or misunderstandings of subtle environmental and internal cues. It seems clear from the above discussion that one of our tasks in stuttering therapy is to help the stutterer understand himself in the interpersonal situation and not distort or confuse the relationship of himself to others.

In other words, it does not seem to be a matter of helping the stutterer to simply recognize outside cues in order to progress in therapy. On the contrary, it seems to be more a matter of helping the stutterer understand his distortions and misunderstandings of his stuttering (as part of himself) as they are related to the interpersonal communicative situation. This would seem to be a more realistic and constructive means of reducing or eliminating the conflict that is responsible for the disturbed communication. Trying to convince stutterers in therapy that they will have to accept stuttering as a part of themselves seems to be a negative and psychotherapeutically unsound frame of reference to work by. It seems apparent that if we are going to effect a meaningful change in the total behavior of the stutterer we will have to concentrate on the elimination of stuttering from the stutterer's self concept through a more positive type of symptomatic psychotherapy.

C. Experimental Research

Experimental studies of stutterers' self concept have not provided us with the necessary data for an understanding of the dynamics of stuttering as related to self concept. The studies reported in the literature seem to indicate that stutterers have a higher self concept than nonstutterers, Zelen, Sheehan and Bugantal (19), that stutterers have a lower self concept, Wallen (17), and that stutterers have no significant difference in self concept from nonstutterers, Fiedler and Wepman (7). Haney (11) using the TAT concluded that stutterers had contradictory concepts of themselves. Fitzpatrick (9) investigated the body image of secondary stutterers as revealed by their self drawings when compared with their drawings of the ideal speaker. This study indicated that, "The self drawing depicted a lesser degree of speech proficiency than did the drawing of the ideal."

Clark (4) commented on Fitzpatrick's study by indicating that, "The body image formed by stutterers must be very difficult to tolerate. Perhaps the relapse or the breakdown of fluency that occurs with stutterers, can be more easily understood in terms of body-image concept. One of the hopes for successful therapy with speech or multiple handicapped individuals is to understand the significance of body-image problems and to help alter them."

Reviewing the studies cited above, it appears that research designed to investigate the relationship between self concept and stuttering produced contradictory, inconclusive results. The fact that each study approached the measuring of self concept with a different frame of reference as well as a different measuring technique seems to have contributed to the uncertainty and contradiction reflected in the results. A different approach to research in this area seems to be needed. Witkin (18) has demonstrated that the mode of field approach, i.e., field dependence (global), field independence (synthetic) is significantly related to the degree of "sophistication of body image" as well as the degree of "sense of separate identity." His evidence shows that the presence or absence of pathology are not directly related to the mode of field approach. The evidence also shows that disturbances and pathology assume different forms among people with relatively analytical and relatively global field approaches.

Witkin (18) points out however, that there is a possibility that pathology may be less likely to develop among "intermediates" than among the extremes of the continuum. The mode of field approach was found to cut across some of the conventional classifications categories on psychopathology (paranoid neurotic etc.) however, it was found to be related to symptomatology and psychodynamics. It seems that Witkin's field approach and the instruments he has developed to test his assumptions might be of considerable value to the researcher interested in self concept and stuttering. This approach not only shows promise in clarifying the relationship of self concept and stuttering but also may be helpful in clarifying the perennial problems of differential diagnosis in the area of communication disorders.

From the review presented in this paper, it appears that future research in the area of stuttering and self concept should consider the following:

1. A need *a priori* for a nonelementalistic system that attempts to relate stuttering theory to self concept (personality theory).

2. A need for a more clearly defined operational definition of self concept as it relates to our present knowledge of stuttering. This can logically lead to better theoretical constructs that associate stuttering relapse to self-alienating phenomena or false fluency.

3. A need for a valid and reliable instrument internally consistent with the theory that can adequately measure the dynamic relationship of the variables.

The future of research in this area as well as in other related areas depends upon the creative imagination of the theoretician and the skillful craftsmanship of the experimenter. This undeveloped area of knowledge looks promising only if current experimentation is synthesized into sound systematic theory and put to the test.

D. SUMMARY

The purpose of this paper was to review the literature pertaining to stuttering and self concept and clarify the current issues pertaining to theory and research in this area. A discussion of self concept as it relates to symptomotology and therapy is presented. The methodological and conceptual limitations of the literature are briefly examined and suggestions for additional research are presented.

REFERENCES

1. ADLER, A. The Practice and Theory of Individual Psychology. London: Routledge & Paul, 1929.
2. ALLPORT, G. Becoming; Basic Considerations for a Psychology of Personality. New Haven: Yale Univ. Press, 1955.
3. BLOODSTEIN, O. A rating scale under which stuttering is reduced or absent. *JSHD,* 1950, **15**.
4. CLARK, R. M. The psychological implications for speech and hearing therapists of the characteristics of infantile thinking and body image. *In* Barbara, D. (*Ed.*), ***Psychological Psychiatric Aspects of Speech and Hearing***. Springfield, Ill.: Thomas, 1960.
5. DORSEY, J., & SEEGERS, W. H. Living Consciously: The Science of Self. Detroit: Wayne State Univ. Press, 1959.
6. DOUGLAS, E., & QUARRINGTON, B. The differentiation of interiorized and exteriorized stuttering. *JSHD,* 1952, **17**, 377-385.
7. FIEDLER, F. E., & WEPMAN, J. M. An exploratory investigation of the self concept of stutterers. *JSHD,* 1951, **16**.
8. FISHER, S., & CLEVELAND, S. E. Body Image and Personality. Princeton: Van Nostrand, 1958.
9. FITZPATRICK, J. A. An investigation of the body image in secondary stutterers revealed through self-drawings. Univ. Denver: Doctoral diss., 1959.
10. FROMM, E. Man for Himself. New York: Rinehardt, 1947.
11. HANEY, H. R. Motives implied by the act of stuttering as revealed by prolonged experimental prolongation. Ph.D. Diss., Univ. Southern California.
12. LEWIN, K. Field Theory in Social Science. New York: Harper, 1951.
13. ROGERS, C. R. Client-Centered Therapy. Boston: Houghton Mifflin, 1951.
14. SHEARER, W. A theoretical consideration of the self concept and body image in stuttering therapy. *ASHA,* 1961, **3**, 10-11.
15. SHEEHAN, J. G. Theory and treatment of stuttering as an approach avoidance conflict. *J. of Psychol.,* 1953, **36**, 27-49.
16. VAN RIPER, C. Symptomatic Therapy for Stutterers. *In* Travis, L. (*Ed.*), ***Handbook of Speech Pathology***. New York: Appleton-Century-Crofts, 1957.
17. WALLEN, V. A Q-technique study of the self-concepts of stutterers and nonstutterers. Ph.D. Diss., Boston Univ., 1959.
18. WITKIN, H. A., *et al.* Psychological Differentiation. New York: Wiley, 1962.
19. ZELEN, S. L., SHEEHAN, J. G., & BUGANTAL, J. F. T. Self perceptions in stuttering. *J. Clin. Psychol.,* 1954, **10**, 70-72.

Rutgers University
Newark 2, New Jersey

[illegible]

The purpose of this paper was to review the literature pertaining to [illegible] and [illegible] from the current [illegible] and [illegible] this area. A [illegible] of self concept [illegible] is presented. The methodological and conceptual limitations of the literature are briefly examined and suggestions for [illegible] are presented.

REFERENCES

1. [illegible] A. The Practice and Theory of Individual Psychology. London: Kegan Paul, Trench & Trubner, [illegible]
2. [illegible]
3. [illegible]
4. [illegible]
5. [illegible]
6. [illegible] The differentiation of [illegible] and [illegible]
7. [illegible]
8. [illegible] Body Image and Personality. [illegible]
9. [illegible]
10. [illegible]
11. [illegible]
12. [illegible]
13. [illegible]
14. [illegible]
15. [illegible]
16. [illegible]
17. [illegible]
18. [illegible]
19. [illegible]

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 313-322.

THE DISADVANTAGED GROUP: A CONCEPT APPLICABLE TO THE HANDICAPPED*[1]

Detroit, Michigan

SIDNEY JORDAN

A. INTRODUCTION

A prevalent practice in the literature concerning the physically handicapped is to compare them with minority groups (1, 5, 11, 12, 14, 15). One such example is the following quote from Rusalem:

> In discussing Lewin's statement that minority groups tend to accept the implicit judgment of those who have status even where the judgment is directed against themselves, Meyerson points out, "This is true of the disabled minority as it is of the racial and cultural minorities." By implication, then, one may transfer some findings from one area to another (14, p. 279).

It is not unusual for apparently similar concepts to be grouped together. This grouping can lead to the generalization of research findings from one group to another that appears similar. The application of concepts from one group to another can: (*a*) implement a reevaluation of previous data and theory; or (*b*) lead to confusion wherein the original concept becomes so fluid that it is no longer useful.

The question here is whether it is profitable to regard the disabled[2] as a minority group. There appears to be adequate basis for regarding the disabled under separate concepts from those applicable to minority groups. Development of a concept specifically applicable to the physically handicapped may further

---

* Received in the Editorial Office on October 30, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The concept of the Disadvantaged Group is derived from lectures given by Dr. Herbert Greenberg at Texas Technological College during the Spring Semester, 1957. Dr. Greenberg's definition is that membership in a Disadvantaged Group brings discrimination to an individual in any one of three areas; political, social, and/or economic. The concept as originally proposed, and its interpretation by the author vary in the number of areas. Dr. Greenberg has not submitted the concept as he proposed it for publication.

[2] The terms disability and physically handicapped are not synonymous as employed in this paper. The distinction, derived from K. W. Hamilton "A disability is a condition of impairment, physical or mental, having an objective aspect that can usually be described by a physician . . . A handicap is the cumulative result of the obstacles which disability interposes between the individual and his maximum functional level."

practical and heuristic ends. The purpose of this paper is to develop such a concept. The term Disadvantaged Group designates the characteristic of this concept.

B. CHARACTERISTICS OF DISADVANTAGED GROUP

The Disadvantaged Group is herein defined as that group composed of individuals marked in their socio-cultural affiliations, socio-economic, or professional trade activities, because of a particular, discernible physiological defect.

The one trait (i.e., blindness, deafness, etc.) that the members of the Disadvantaged Group have in common, influences the psychological and social orientations of its members, and is the means by which they are identified. As defined, the concept is applicable on the following levels:

1. *Not Self-Perpetuating*

A Disadvantaged Group is not self-perpetuating. The disability around which the group is built may occur at birth or adventitiously, but it is never perpetuated by the group as an ethic. Limited by their disability from full participation or acceptance within their primary group, the physically handicapped find that their joining a Disadvantaged Group is socially reinforced. The Disadvantaged Group's orientation, in turn, almost assures acceptance, and offers the physically handicapped person an opportunity to attain status and other role relationships otherwise denied him.

The minority group, on the other hand, perpetuates its values and norms through the family unit. Homogeneity is fostered and multi-leveled association with dissimilar groups is not condoned. Unlike a minority group, the very nature of a Disadvantaged Group demands that it rely upon outside sources to foster group cohesion and attain group goals.

2. *Misunderstood Social Reactions*

The social reactions experienced by members of a Disadvantaged Group stem from two basic factors; (*a*) the lack of comprehensibility by an individual not possessing the trait to understand the consequences and limitations of the disability; and (*b*) the handicapping character of the physical defect is thought to diffuse throughout the total personality and behavior of the individual, so that he can be both classified and identified by this one discernible fact.

Using the standards of their own group's affiliations to act as guides, the dominant group can interpret, or misinterpret, minority group activities. The differences claimed serve a number of purposes, such as proof of the dominant group's superiority, a focus for the projection of unacceptable wishes, and a means of perpetuating social distance, etc. The same standards do not apply to the physically handicapped. To know what it is like to have a sense is not an adequate guide for interpreting what it is like to lack a sense modality. Unable to fully understand the limitations imposed, or the attainments possible, the non-handicapped rely upon stereotypes of the particular physical handicap to characterize the individual or the group.

The stereotypes, which vary to some degree for each physical handicap (1, 2, 5, 6, 8, 9, 10, 11, 14), reflect and reinforce culturally derived attitudes. Whether it is the belief that the lack of a sense modality forces the person to live in a void, or that any attainment by a physically handicapped person is to be considered miraculous, the reaction displayed is to the handicap, and not to the individual.

3. *Imposed Psychological Distance*

Members of a Disadvantaged Group must cope with a psychological distance that is imposed upon them, and which they do not directly foster. Factors that indicate that the psychological distance is imposed, are: (*a*) the obvious emotional overtones of the behaviors displayed by the non-handicapped to the physically handicapped; and (*b*) the desire that the disabled be isolated "for their own good."

The inability to know what it is like to lack a sense modality does not totally explain the psychological reactions experienced by the disadvantaged in their dealings with the non-handicapped. Physically handicapped persons become accustomed to experiencing excessively emotional reactions in most of the situations where interaction between the handicapped and non-handicapped take place (2, 3, 4, 5, 6, 14). Overcoming this reaction demands that the physically handicapped person become known as an individual in spite of his disability. The opportunity to do this is limited by social factors that limit the degree and type of interaction possible, and the personalities of both persons. The most overt means of maintaining psychological distance and avoiding interpersonal contact, is through segregation. While a study done by Lukoff and Whiteman is specifically concerned with attitudes toward the blind, one of their findings is applicable to this discussion. They state:

> There is a positive relationship between tendencies to feel pity for blind people on the one hand, and the tendency to espouse community "segregation" for the blind on the other. Our data show that this is not due to "pitiers," seeing the blind as inferior, but to the fact that those

> that pity the blind feel that it is right that they should be given special protection and help, which is represented at the community level by a "segregated" service, and recreational pattern. Thus, espousal of segregation for blind persons has quite different meaning than espousal for ethnic groups (10, p. 42).

The segregation of minority or ethnic groups takes different forms and serves different purposes than the segregation of the physically handicapped. The segregation of a minority group is enforced by external pressures and the desire of the group itself. By limiting the interaction of the group's members with others, the minority group can maintain a homogeneity and perpetuate its values and norms. Time, however, the influence of educating the young in schools oriented toward majority group standards, and the rewards offered for conforming to the majority group's social demands, lead to breakdowns in minority group affiliations and values. The disadvantaged, on the other hand, are not segregated because of their minority, or ethnic group backgrounds, and they do not desire the psychological distance imposed. It is the implications of the handicap, as interpreted by the non-handicapped, that precludes the acceptance of the individual and limits his ability to interact.

4. *Outside Help Sanctioned*

Behaviours designed to "help the physically handicapped" are positively sanctioned. Antipathetic attitudes, on the other hand, are not sanctioned, and when displayed, produce negative reactions toward the source from the non-handicapped. This orientation limits interpersonal relations with the disabled in two ways: (*a*) it limits the testing out of what is right and wrong behavior toward a disabled person for fear of negative reactions from the non-handicapped environment; and (*b*) culturally prescribed santions tend to preclude behaviors that would produce a release of tension and enable physically handicapped persons to fully participate in non-handicapped group structures.

The non-handicapped person, when dealing with the physically handicapped, must cope with limitations (*a*) and (*b*) listed above. Because of these limitations it is easier to rely upon preconceived notions of what the disabled are thought to be like, than to test the suppositions' veracity at the price of misunderstanding or scorn. Social stereotypes and the "definition" of the limitations caused by the disability become the basis for an "objective appraisal" of the handicapped individual. The objective appraisal, in turn, becomes the basis for a number of prescribed behavior patterns in interpersonal situations. These behavior patterns take the form of an over-concern about, and a compulsive need to immediately satisfy the disabled person's wants, and the avoidance of any discussion or activity in which he cannot knowingly participate, or would be at a disadvantage, if he should participate. The handicapped person's companion, if available, is often relied upon to help structure the situation. While directing the conversation to the mediary, the wants, personality characteristics, and likes of the disabled person can be discussed as if he or she were not there. The psychological distance that is maintained by these behavior patterns is not recognized, and if brought to the non-handicapped person's attention, would be rationalized, and not thought of as unaccepting.

The overt expression of negative attitudes towards members of a minority group, on the other hand, is socially acceptable by many of the groups making up the majority. Negative attitudes about minority groups facilitate their being used as scapegoats, and are employed as the basis for manipulating the group's members economically and socially, with no misgivings. Their "peculiarities" allow joking about them in a flippant fashion, or paying them at different wage scales than members of the majority group. As long as a minority group member holds a subservient position, interpersonal relations are readily attained and are easily maintained. Confidences and jocularities, the imputation of moral freedoms, physiological capabilities, or innate shrewdness, are rendered freely, or are encouraged. Obviously, none of these attitudes would be allowed free expression in dealing with members of Disadvantaged Groups.

5. *Dependence Upon External Agents*

The lack, or restricted use, of a sense modality, causes the individual to be unable to function in some areas and fosters a dependency upon external agents to help compensate for the limitations imposed. The societal desire to help the physically handicapped takes two forms: (*a*) the unorganized and mostly individual attempts to help a specific disabled person in a specific situation; and (*b*) the institutional attitudes, as reflected in legal statutes and service agencies, where organized attempts are made to help the physically handicapped as members of a Disadvantaged Group. The societal characteristic of both reactions can be a source of secondary gain for the physically handicapped.

When dealing with a disabled person on an individual level, the desire to be the outside agent helping him serves an objective and socially espoused purpose. The degree and type of help usually given, however, is indicative of the confusion felt by not knowing how to act toward, or what is desired, by the handicapped person. Examples of the excessive reactions, more than occasionally displayed, usually are seen when a cripple attempts to carry a parcel, and it is taken away for him without asking, in spite of, or in ignorance of, his wishes or capabilities. For a blind person to find his destination by himself is considered impossible or miraculous. He, therefore, is to be helped to his destination, whether he wants the help or not. Hearing aids improve the audition of persons with impaired hearing. Therefore, if the person doesn't have a hearing aid, get him one. The fact that there is no attempt made to understand the physically handicapped person as an individual, with feelings and judgments of his own, is strongly evident. The help is thrust upon him. Whether he wants or can use what is being offered, is virtually a secondary concern.

Examples of institutionalized attitudes are evident on the national, state, and community levels. They include the division of welfare categories, so that recipients of Aid to the Blind, and Aid to the Disabled, receive built-in considerations because of their handicap. Special tax supported agencies are maintained to educate, train, and rehabilitate group members. A number of disabled groups receive special tax considerations and differentiated treatment in the courts (2).

Private agencies act as a direct means by which specific physically handicapped groups can be helped. These agencies collect funds by focusing the attention of the giver on one or more of the limitations caused by the disability, and upon rehabilitation. They do not distribute funds as governmental agencies do. Their centers of interest may extend from the offering of the educational or diagnostic facilities that may prevent an incipient loss from becoming severe, to the supplying of specific compensatory aids, counseling, or recreational services, and the maintenance of sheltered workshops for those already disabled.

The handicapped individual can experience a degree of secondary gain by manipulating the institutionalized and individual help thrust upon him. If properly utilized, the individual need not assert himself, or achieve, in competition with non-handicapped persons, in order to survive. He need not work, but if he desires to work, a number of prescribed jobs are set aside for him. If the disabled individual is involved in a social disturbance, the blame is either put onto others, or is attributed to the handicap, not to the individual as an individual. Minor breaks in social mores are either overlooked, or if punished, receive differential treatment from the law.

The secondary gains allowed members of a minority group take different and less compensatory form than those provided members of Disadvantaged

Groups. As long as the individual maintains his identification with the minority group, he knows what is behaviorally expected and tolerated. The group affords him social and recreational releases, wherein he can attain a level of status not possible in reactions with the majority group. The group also acts as a buffer against inadequacy, by attributing failure to the prejudices of the majority group. Conformation of the prejudicial attitudes are found in the limited job offerings for minority group members, and the differential and negative treatment received from judicial and law enforcement agencies.

6. *External Definition of Adjustment*

It is not the physically handicapped person who defines whether he is socially or psychologically adjusted. The judgment is made of him by the non-handicapped, or representatives of agencies, whose policy is structured by the social-psychological interpretations of what Disadvantaged Group members are supposed to be like. The reaction of handicapped persons to these interpretations extend along a continuum. At one end is the individual who submits to the "imputation of inferiority"[3] associated with his disability, and accepts the segregation and secondary gains afforded. This individual is assured a subsistence as a welfare recipient, or earns his livelihood in a number of prescribed areas, or in a sheltered placement. While these jobs pay comparatively little, they present the physically handicapped person with an opportunity to attain a higher status, or a significant role within a structured environment. Individuals at this end of the continuum have adjusted, in the societal sense.

On the other end of the continuum is the physically handicapped individual who does not accept the psycho-social implications of his disability. Desiring to test his capabilities as an equal in a competitive setting, he relies upon the devices designed to help him compensate for his disability, and tries to deny his dependency upon others. He does not employ, to any marked degree, the secondary gains allowed him, and overtly refuses to conform to majority expectations. The individual aspiring to this goal is not considered to have adjusted to his physical handicap in the psycho-social sense.

While the handicapped person may desire to be accepted as an equal in a competitive setting, he is rarely given the opportunity to do so. Factors beyond the individual's control effectively limit his getting a job equal to his skills or training in a setting not oriented to his particular disability. Among the obstructions are the attitudes of management, and the antipathy of the non-handicapped workers with whom the disabled worker would be employed (13). If the disabled person is given an opportunity to work as an equal in an integrated setting, he becomes a competitive force, and the attitudes that usually define his status and role do not prove applicable. The behavior that would permit a change in attitudes and expectations is limited by the handicapped worker's physical impairment, and the able-bodied workers' inability to accept as equal, or as supervisor, a person who cannot deny his dependency upon others, and outside devices, to perform the same tasks that the non-handicapped can do without guides. During times of employment crisis, the non-physically handicapped worker tends to feel threatened by the adequate performance of those he perceives as inferior. He attributes his own deficiencies to the difficulties caused by working with a disabled person. Management, relying upon their own culturally engendered, preconceived notions, frequently take the position that disabled workers demand special considerations, and costs, and impair the morale and performance of other workers. This opinion is maintained, despite objective criteria as to the adequacy of performance and production by the physically handicapped.

---

[3] The phrase, derived from Chevigny and Braverman (5) is employed by the authors specifically for the blind. It is employed in this context because it is applicable to all Disadvantaged Groups, as defined.

Just as the sources of conflict differ for the minority group and Disadvantaged Group members, there are differences in the social interpretations of the individual's adjustment to being a member to either group. Unlike the disadvantaged, minority group affiliation is not constantly reinforced by the majority. While many members of the majority may desire that the minority group member know, and keep his place, subtle and coercive factors, such as interpersonal contact, education, the mass media, legal pressures, and the desire for economic or occupational advancement, foster identification with the dominant group's values and orientations. The same factors that lead to the identification with the dominant group thwart the minority group's attempts to segregate itself, in order to maintain its own ethic.

Should the minority group member continue to adhere to his group's values, there is still a degree of tacit acceptance in a competitive setting, as long as the competition is open, and follows the rules set down by the dominant group. Acceptance in employment or political situations does not imply acceptability on a social level. The inability to socialize with members of a dominant group can be a major source of frustration for a minority group member.

The minority group can be perceived as a group in transition. While constantly being influenced by subtle and coercive factors to assimilate, it is also able to assert itself in a competitive setting. A Disadvantaged Group, on the other hand, finds itself in a "social limbo."[4] Segregation is enforced and attempted assimilation is frustrated, making it difficult to function as an effective social influence.

C. SUMMARY AND CONCLUSION

With the ever increasing number of studies of the physically handicapped, there is a need for a rigorous examination of basic premises and conceptualizations. One widely employed premise is the equation of physically handicapped and minority groups. It is held that adherence to this premise leads to conceptual confusion, and impedes heuristic goals, for the facts derived from the study of the physically handicapped are non-identical with, or are contrary to, the findings of studies of minority groups. This paper is an attempt to develop a concept that is specifically applicable to the physically handicapped. The phrase used to designate the concept is Disadvantaged Group. To demonstrate the dissimilarities of minority and Disadvantaged Groups, comparisons were drawn on six inferences, derived from the basic premise underlying the proposed concept.

REFERENCES

1. BARKER, R. G. The social psychology of physical disability. *J. Soc. Issues*, 1948, **4**, 28-38.
2. BARKER, R. G., WRIGHT, B. A., MYERSON, L., & GONICK, M. R. Adjustment to Physical Handicap and Illness: A Survey of the Social Psychology of Physique and Disability. New York: Social Science Research Council Bulletin 55, 1953.
3. BEST, H. Deafness and the Deaf in the United States. New York: Macmillan, 1943.
4. CARROLL, REV. T. J. Blindness: What It Is, What It Does, and How to Live With It. Boston: Little, Brown and Co., 1961.
5. CHEVIGNY, H., & BRAVERMAN, S. The Adjustment of the Blind. New Haven: Yale Univ. Press, 1950.
6. GOWMAN, A. G. The War Blind In American Social Structure. New York: American Foundation for the Blind, 1957.
7. HAMILTON, K. W. Counseling the Handicapped in the Rehabilitation Process. New York: Ronald, 1950.
8. KESSLER, H. H. Rehabilitation of the Physically Handicapped. New York: Columbia Univ. Press, 1947.
9. LEVINE, E. S. The Psychology of Deafness. New York: Columbia Univ. Press, 1960.
10. LUKOFF, I. F., & WHITEMAN, M. Attitudes toward blindness, *The New Outlook for the Blind,* 1961, **55**, 39-44.

---

[4] A. G. Gowman (6) employs this phrase to describe the psycho-social reactions experienced by the blind. It is held that the phrase is also applicable to other Disadvantaged Groups.

11. MUSSEN, P. H., & BARKER, R. G. Attitudes toward cripples. *J. Abn. & Soc. Psychol.*, 1944, **39**, 351-355.

12. MYERSON, L. Physical disability as a social psychological problem. *J. Soc. Issues*, 1948, **4**, 2-10.

13. NOLAND, E. W., & BAKKE, E. W. Workers Wanted: A Study of Employers' Hiring Policies, Preferences and Practices in New Haven and Charlotte. New York: Harper, 1949.

14. RUSALEM, H. The environmental supports of public attitudes toward the blind. *The New Outlook for the Blind*, 1950, **44**, 277-288.

15. WRIGHT, B. A. Physical Disability—A Psychological Approach. New York: Harper, 1960.

18202 Stoepel
Detroit 21, Michigan

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 323-328.

LATERAL DOMINANCE CHARACTERISTICS AND READING ACHIEVEMENT IN THE FIRST GRADE*

Division of Social Sciences, University of California, Riverside

IRVING H. BALOW

A. THE PROBLEM

The literature considering lateral dominance and its relationship to reading is very extensive. Two major points of view exist concerning this relationship. One of these is the result of the work of Dearborn (2) and Monroe (4) both of whom believed that left-eye dominance, lack of dominance, or crossed dominance were conditions associated with reading disability. A more recent study by Harris (3) represents a second point of view that directional confusion is an important factor in reading disability. He found mixed-hand preference and confusion in identifying right and left significantly related to reading disability.

Harris suggests that one reason for significant results in his study when other studies have failed to find differences can be found in the dominance tests used—the tests in his study being sensitive indicators of directional confusion.

The studies of Smith (5) and of Clark (1) lend some credence to Harris' claim. Smith, using several tests of hand preference, found that only the Van Riper Angle-Board Test, a sensitive test of hand dominance, differentiated between good and poor readers. Clark found that all children who had left-hand tendencies scored significantly lower on an achievement test battery than did the children without such tendencies (1).

The problem considered in this study was whether the dominance anomalies specified by these two points of view are significantly associated with reading achievement in the first grade.

B. DESIGN

This study was carried out in a middle to upper-middle class suburb of St. Paul, Minnesota. Thirteen first grade classrooms were randomly selected and the children in these classes constituted the sample.

The 320 children in these classes at the beginning of the school year were

---

* Received in the Editorial Office on November 13, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

individually tested with the Harris Tests of Lateral Dominance by the investigator. These tests were completed before the end of the third week of school. The Gates Reading Readiness Tests were administered during the second and third weeks of school by the investigator with the assistance of the classroom teacher. In December the Lorge-Thorndike Intelligence Test was administered, and in the last week of February the Gates Primary Reading Tests, PWR-Word Recognition and PPR-Paragraph Reading.

Of the 320 children originally tested, 18 left the district before the February testing, thus, the true sample consisted of 302 children—151 boys and 151 girls.

The groups formed to test each hypothesis were first tested for equality of IQ scores. In each case the mean IQ of the groups involved were statistically equal.

C. RESULTS

The 302 children were classified according to hand preference (right and mixed or left) and eye preference (strong right, mixed, and strong left) in order to test the first hypothesis that there are no differences in reading readiness score, word recognition score, or paragraph reading score associated with hand dominance or eye dominance or an interaction of these factors. It was not possible to use a separate classification for the left-handed as only fifteen children in the sample had left-hand tendencies according to the dominance tests. This two-way classification resulted in six cells with at least twenty-eight scores in each cell. Twenty-seven subjects were randomly selected to represent each cell. Table 1 shows the mean scores for reading readiness, word recognition, and paragraph reading for the twenty-seven scores in each cell.

Each row of mean scores in Table 1 was tested by the analysis of variance to determine if the main effects of hand preference and eye preference or their interaction was significant.

The probability of the *F*-ratios, in each case, was greater than .05. The first null hypothesis was therefore accepted. Hand preference and eye preference, singly or in interaction, were not significantly associated with reading readiness, word reading, or paragraph reading score. For this group of first grade children, hand preference and/or eye preference was not related to reading achievement.

To test the second hypothesis, the 302 pupils were classified according to dominance (normal—those having hand and eye preference on the same side of the body, crossed—those having hand preference on one side and

TABLE 1

MEAN SCORES FOR 162 PUPILS CLASSIFIED BY HAND AND EYE PREFERENCE

	Right handed			Mixed or left handed		
	Strong right eye	Mixed eye	Strong left eye	Strong right eye	Mixed eye	Strong left eye
Reading readiness	63.037	61.148	61.555	62.037	62.444	62.555
Word reading	25.667	23.518	25.629	22.222	24.185	20.925
Paragraph reading	12.851	11.481	12.333	12.222	12.851	11.222

eye preference on the other side of the body, and mixed—those having mixed hand preference, according to the Harris Tests of Lateral Dominance), and knowledge of right and left (normal, hesitant, and confused). This two-way classification resulted in nine cells. Fifteen subjects were randomly selected from those classified in each cell and their scores used in the analysis. The mean scores resulting from this classification are shown in Table 2.

The differences between the mean scores in each row of Table 2 were tested for significance using the analysis of variance.

Each of the *F*-ratios had a probability greater than .05. The hypothesis of no significant differences in reading readiness, word reading, or paragraph reading associated with normal, crossed, or mixed dominance; knowledge of right and left; or an interaction of these factors was accepted. For these children reading achievement was not significantly related to normal, crossed, or mixed dominance, or to their knowledge of right and left.

The third hypothesis of no significant differences in reading readiness, word reading, or paragraph reading associated with strong, moderate, or mixed hand dominance; knowledge of right and left, or an interaction of those factors, was stated to provide a direct check on Harris' findings using first grade children. The 302 children were classified as strong, moderate, or mixed in hand dominance and knowledge of right and left. Seventeen subjects were randomly selected from those classified in each cell and their scores used in the analysis. The mean scores resulting from this classification are shown in Table 3.

The differences between the mean scores in each row of Table 3 were tested for significance using the analysis of variance.

Each of the *F*-ratios had a probability greater than .05. The hypothesis of no significant differences in reading readiness, word reading, or paragraph reading associated with strength of hand dominance, knowledge of right and left, or an interaction of these factors was accepted.

D. CONCLUSIONS

1. The first grade child's hand dominance, or lack of dominance, does not seem to be related to reading achievement. The children with mixed hand dominance, as a group, scored as high as the children with consistent hand dominance on each measure of reading achievement. The same was true for eye dominance for these children.

2. Having the dominant hand and eye on the same side of the body, on opposite sides of the body, or having mixed hand dominance has no significant effect on reading achievement in the first grade.

TABLE 2

MEAN SCORES OF 135 CHILDREN CLASSIFIED ACCORDING TO LATERAL DOMINANCE AND KNOWLEDGE OF RIGHT AND LEFT

Knowledge of right and left	Normal dominance			Crossed dominance			Mixed dominance		
	Normal	Hesitant	Confused	Normal	Hesitant	Confused	Normal	Hesitant	Confused
Reading readiness	69.066	61.466	62.533	61.600	58.400	61.266	72.000	59.066	62.800
Word reading	27.733	24.600	25.133	27.000	26.466	24.800	29.000	21.266	21.466
Paragraph reading	13.333	12.733	13.866	13.600	12.466	11.800	14.800	10.466	11.600

TABLE 3

MEAN SCORES OF 153 CHILDREN CLASSIFIED ACCORDING TO STRENGTH OF HAND DOMINANCE AND KNOWLEDGE OF RIGHT AND LEFT
(17 scores per cell)

Hand dominance	Strong			Moderate			Mixed		
Knowledge of right and left	Normal	Hesitant	Confused	Normal	Hesitant	Confused	Normal	Hesitant	Confused
Reading readiness	69.588	53.352	60.352	63.176	61.294	60.176	63.000	61.058	62.294
Word reading	26.764	21.647	25.352	26.764	28.764	18.352	22.411	24.000	21.470
Paragraph reading	12.529	10.647	14.176	12.764	13.235	10.705	12.235	11.823	11.941

3. In this sample of six-year-olds, 42.4 per cent manifested mixed hand preference. In Harris' group (3) only 18 per cent of seven-year-olds were classified as having mixed hand dominance. This difference indicates that a large number of children establish a consistent directional preference during the first-grade. It is hypothesized that those children with mixed hand dominance who learn to read well in first grade establish consistent dominance during the year and that this consistency is at least partially the result of the directional knowledge learned from successful reading instruction. A follow-up study of these children is planned for their second-grade year to determine if this hypothesis is tenable.

4. Knowledge of right and left is not a significant factor in reading achievement in the first grade. The children who were confused about right and left at the beginning of the year achieved as much in reading as the children who knew right and left.

5. Strength of hand dominance is not a significant factor in reading achievement in the first grade.

6. Among the classifications used in this study, there is no combination of hand dominance and eye dominance, hand-eye dominance and knowledge of right and left, or strength of dominance and knowledge of right and left significantly associated with reading achievement.

7. Lateral dominance screening of children at the beginning of the first grade will not provide the teacher with information which will help her spot those children who are likely to have difficulty in learning to read.

REFERENCES

1. CLARK, M. M. Left Handedness. London: Univ. London Press, 1957.
2. DEARBORN, W. F. Structural factors which condition special disability in reading. *Proceedings of the American Association for Mental Deficiency*, 1933, **38**, 266-283.
3. HARRIS, A. J. Lateral dominance, directional confusion, and reading disability. *J. of Psychol.*, 1957, **44**, 283-294.
4. MONROE, M. Children Who Cannot Read. Chicago: Univ. Chicago Press, 1932.
5. SMITH, L. C. A study of laterality characteristics of retarded readers and reading achievers. Ph.D. thesis, Temple Univ., 1949, University Microfilms, Ann Arbor, Michigan.

Division of Social Sciences
University of California
Riverside, California

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 329-336.

FAMILY CONSTELLATION AND SCHIZOPHRENIA*[1]

Agnews State Hospital, San Jose, California

ARTHUR BURTON AND JOSEPH W. BIRD

A. INTRODUCTION AND PROBLEM

That the dynamics of the family situation are related to the psychopathology of any of its members, was proposed even before Freud formalized it in such conceptions as the Oedipal complex, and in others. While family interaction as a possible causal determinant in the neuroses and psychoses has waxed and waned over the years, we now see a resurgence of interest in the matter of schizophrenia (3, 9).

One family variable which has come in for considerable scrutiny by investigators interested in the areas of personality theory, developmental psychology, and psychopathology has been birth order. Probably the first significant writings relevant to the issue of birth order and personality were those of Adler (1, pp. 144-154). In line with his interest in the social determinants of personality, Adler observed that the personalities of the oldest, middle, and youngest child in a family were likely to be quite different. He attributed these differences to the distinctive experiences of each child as a member of the social group. Neurotics, criminals, drunkards, and perverts, he observed, are often first-born children. He also observed that, next to the oldest child, the youngest child in the family is most likely to become a problem child and a neurotic, maladjusted adult.

Adler saw the second or middle child as ambitious and as constantly striving to surpass his older sibling. Although tending to be rebellious and envious, he is likely to be better adjusted than either his older or younger sibling.

Adler's clinical conjectures about birth order have been tested many times, but most of the findings have failed to lend support to them (8). Of the studies which do tend to support Adler's views (2), the results have not been unequivocal nor have the studies always evidenced sufficient control of the

---

* Received in the Editorial Office on November 19, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported by a research grant from the California Department of Mental Hygiene. The authors wish to gratefully acknowledge the assistance of William Cobourn and Donald L. Lunde in the collection of the data. We are also indebted to Dr. Don D. Jackson for his encouragement and assistance.

relevant variables. It should be noted that the earlier writers tended to assume that families and position in families were static rather than dynamic.

Birth order in schizophrenia has also come in for some attention. Grosz and Miller (4), in reporting one of the better designed and controlled studies, point out that previous investigations had failed to control for size of family; that is, large families were grouped with small families and incomplete families were merged with completed ones. In their study, only patients from families with three children were employed. They further restricted their sample to schizophrenic patients with two normal siblings; twins and adopted children were excluded. Furthermore, the sample included only patients from families in which the siblings were raised together in the same family constellation throughout their childhood. With a sample of 156 schizophrenics, they did not find a significant difference in birth order.

Using manic-depressive patients as controls, essentially similar findings were obtained by Pollock *et. al.* (12) in another study of birth order in schizophrenia.

In studying the case histories of a large number of patients admitted to a state hospital, Wahl (13) concluded that his data did not lend support to a hypothesis of increased vulnerability to schizophrenia with any particular birth placement. As with other studies in this area, Wahl reported his findings in terms of percentages, and there was apparently no control for variability in family size.

Much less is known regarding the influence of the specific gender of the child for the development of schizophrenic pathology. While concluding that the many studies on ordinal position reveal no significance, Jackson (7) does claim to have evidence that the sex of the schizophrenic makes a difference as to the importance of the pathogenic parent. Lidz and his group (9, 10, 11) would seem to support this contention.

That the gender of the child influences character development is now well established, but we are not on such safe ground when it comes to schizophrenia. Males and females come in unequal numbers as schizophrenic patients to state hospitals, but there may be artifacts of admission, retention, or diagnosis which account for the larger number of males in the hospital population.

In her observations of schizophrenics in a large state hospital, Hilgard (5) noted that families from which these patients came tended to be larger. Jackson (7) has concurred in this observation. Wahl (13), similarly, found a greater fecundity in families containing a schizophrenic patient. For both males and females the number of children per family in his study averaged

4.1, but at that time insurance data revealed the mean fecundity to be 2.2 offspring. Also, among his 231 male patients, 52 per cent came from families of four or more children.

If it is possibly true that the genesis of schizophrenia is more closely related to the family matrix than we have heretofore known, then more research effort on schizophrenia should be concentrated on the families of such patients. Thus, such conceptions as "schizophrenogenic mother" or "schizophrenogenic family" need experimental or empirical validation before their putatively clinical usefulness can be corroborated. Unfortunately, it is rarely possible to be present at the creation of a schizophrenic patient, and we rely most often on treatment settings for speculations on past genesis. Nor does the fact that family therapy may occasionally change the family pattern, and thus relieve the patient, alter the *ad hoc* nature of such assumptions regarding the family. The validity of such assumptions and inferences remains to be determined.

Large-sample statistical studies of the life history material of schizophrenic patients have, to some extent, been disappointing. Partially, this may be due to the fact that the use of large samples has led some investigators to assume that no other controls were necessary. Thus, such data frequently represents a pot pourri of variables and observers indistinguishable even in the breech. More serious yet is the frequent omission of any control group whatsoever. The conclusions drawn from Wahl's (13) data raise serious questions due to the lack of such control.

Despite the shortcomings of the life history approach, and in the absence of more precise measurement techniques in research on schizophrenia, large scale statistical studies may make substantial contributions to our knowledge of the disorder, and it is possible to minimize such deficiencies as incomplete data, retrospective falsification, interobserver variability, culturally biased samplings, etc. It must also be recalled that large samples make some generalizations possible in an area notorious for formulations based upon one, two, or three cases.

The present study is an attempt to investigate birth order and gender of hospitalized schizophrenic patients and to compare these findings with similar data on hospitalized nonschizophrenics. The study further compared differences in family size of the two patient groups.

B. PROCEDURE

The schizophrenic ($N = 241$) and nonschizophrenic ($N = 286$) groups constituted the patients consecutively admitted to a state hospital between September and December, 1961.

Life history material on all patients in these groups was systematically examined for family data including sex, ordinal position, size of family, and a number of other items. Incomplete or internally unreliable data were discarded. Principal sources of the data were the social work anamnesis, the hospital admission form, and the psychiatric examination report—usually all three. In assigning patients to groups, the official hospital diagnosis was employed.[2]

The control group was comprised of nonschizophrenic patients with diagnoses of various functional disorders. These included affective psychotics, other nonorganic psychotics, chronic alcoholics, severe psychoneurotics, severe character disorders, etc.

There is sufficient evidence in the literature to indicate that state hospital diagnoses are considerably suspect and some diagnostic overlap undoubtedly exists between our two groups. This is felt, however, to be minimal in face of the precautions taken. Patients with proven or suspected organic or neurological disorders were excluded from the study. The material was recorded by two investigators on the basis of previously agreed criteria.

1. *Samples*

During the period of the study, 241 of the 527 patients admitted with functional disorders, were diagnosed as schizophrenic, while 286 were nonschizophrenic.

As has been repeatedly observed, the schizophrenic patient is, on the average, younger in age at first hospitalization than the nonschizophrenic. In the present sample, the age at first hospitalization of the schizophrenic patients ranged from 18 to 67 years, with a median age of 33 years. The nonschizophrenic group, while showing a comparable range in age (19 to 73 years), had a median age of 44 years at the time of their first admission.

In the schizophrenic group there were slightly more females than males (129 and 112 respectively). The order was reversed in the nonschizophrenic group with males outnumbering females 157 to 129.

In demographic studies in state hospitals, it is important to consider the characteristics of the cultural and socioeconomic populations in the area from which the patients are drawn. The hospital in which the present study was conducted is situated within the limits of a moderate size (approximately 200,000 population) city on the West coast. In a recent study at the same hospital, Hilgard and Newman (6) pointed out that "cases admitted to the state hospital seemingly no longer differ much in educational and occupational status from the population at large." They found the median educational level to be high school graduation, and 21 per cent of the patients had some college education. This is above the general level of the white urban population 25 years and over as reported by the U.S. Census. Because the level in the urban West is somewhat higher than the U.S. as a whole, they felt it was a reasonable conjecture that the hospital population represents a fair cross section of educational levels in the community.

---

[2] In the usual state hospital, a number of diagnoses are made at various points in the patient's treatment program and which are recorded in his case file. Thus, the admission psychiatrist, the psychiatrist who prepares the mental examination, and the ward doctor all describe the patient diagnostically. The "official" diagnosis is an independent collation of all previous observations and becomes a matter of central office record. For this reason, only the latter diagnosis was used in the present study.

In terms of occupational distribution, their results appeared to confirm the educational one in showing the generally representative level of the admission population.

Similarly, in our study, the distributions of religious preferences were found to approximate the denominational distributions in the area. Forty-four per cent of the schizophrenic group and 40 per cent of the nonschizophrenic group were Catholic, while Protestants constituted 54 and 58 per cent of the respective groups. Jewish patients accounted for 2 per cent in each group.

C. RESULTS

During the period of the study, 527 patients on whom the criterion anamnestic material was obtained were admitted to the hospital. Table 1 shows the distributions by sex and diagnosis of these patients.

TABLE 1
DISTRIBUTIONS OF SEX AND DIAGNOSIS OF PATIENT SAMPLES

Sex	Schizophrenics	Nonschizophrenics
Males	112	157
Females	129	129
Total	241	286

While comparative data were not obtained, it is felt that the diagnostic distribution is representative of the year-by-year admissions to the hospital. The schizophrenic group accounted for 45.7 per cent of the total sample, while the nonschizophrenic patients represented 54.3 per cent.

Although, as noted above, previous demographic studies have found male schizophrenic patients outnumbering females in admissions to state hospitals, the differences, although in the opposite direction, were negligible. Females outnumbered males in the schizophrenic sample; the order was reversed in the nonschizophrenic group.

The observed differences in sex distribution by diagnosis were tested by chi square. The tabulated and calculated chi-square values are presented in Table 2.

The differences were not found to be significant and the null hypothesis could not, therefore, be rejected ($p > .05$).

TABLE 2
DIFFERENTIAL DISTRIBUTIONS OF SEX, BIRTH ORDER, AND FAMILY SIZE IN SCHIZOPHRENICS AND NONSCHIZOPHRENICS

	N	Tabulated χ^2 ($p = .05$)	Calculated χ^2	p
Sex	527	3.84	3.71	$> .05$
Birth order	464	5.99	2.48	$> .05$
Number of children	527	14.07	12.71	$> .05$

No attempt was made to compare the family size from which our patient groups came with any nonhospitalized population. The investigators feel that use of any national norms, e.g., U.S. Census data, would suffer from serious confounding of the results by reason of the extreme variability resulting from the grouping of very heterogeneous ethnic, socioeconomic, cultural, and regional populations. The mean number of children in the family (including the patient) for the two groups is presented in Table 3.

TABLE 3
MEAN NUMBER OF CHILDREN IN FAMILIES OF SCHIZOPHRENICS AND NONSCHIZOPHRENICS

	Number of children in family, including patient			
	Schizophrenics		Nonschizophrenics	
	Mean	σ	Mean	σ
Males	3.84	2.02	4.14	2.03
Females	3.76	2.14	3.50	2.12
Total	3.85	1.99	3.85	1.28

These data have been further broken down by sex of the patient for the two groups. It may be noted that the results are strikingly similar for the two groups, with the total means being identical. The chi-square comparison failed to reach significance ($p > .05$). (See Table 3.)

In comparing the groups on differences in ordinal position in the family, the patients were divided into first-born, last-born, and those born other than first or last in order. Only-children, twins, and adopted children or those with adopted siblings were excluded from this comparison. The chi-square values are presented in Table 2. The observed differences were not found II to be significant and the null hypothesis could not, therefore, be rejected ($p > .05$).

D. Discussion

There is no a priori reason to believe that parents who produce a schizophrenic patient should have more progeny than those who produce a nonschizophrenic patient, and, indeed, this is the case. Apparently there is nothing specific to progenitors of schizophrenia which would lead them to have greater fertility than, say, the progenitors of a manic-depressive or hospitalized neurotic. Of course, we are dealing here with highly complex biological and psychological phenomena which are not to be interpreted in such a causally simplified way. The temptation is also great to relate the findings of 3.85 progeny of the parents of schizophrenics to past census data of several decades, when our patients were conceived. Because of our belief in the inapplicability of census data in the present instance, and because of the large standard error (± 1.99) which accompanies the figure above, we are refraining from such comparison. The conclusion, then, from our data would be that whatever factors operate within a family in terms of the numbers of siblings present do not as such predispose a person to schizophrenia as a form of mental illness. The answer to such question must be sought in the *qualia* of the family experience rather than the absolute number of its members.

It is easy to accept at face value that the pull upon any specific child by a specific parent will vary whether they are like-sexed or unlike-sexed. While the Oedipal situation is only one among several developmental stresses in the child's development, it is an important one, and one for which failure in mastery can lead to severe character distortion. There is in our opinion a considerable dynamism in the fact that a mother and male child complement each other biologically and psychologically. A mother and female child, however, neutralize or oppose each other, in this sense, so that an essential symbiosis is lost in their relationship. Despite the wide acceptance of such a psychoanalytic formulation our findings fail to show that the sex of the child who becomes schizophrenic has anything to do with the specificity of his illness. Thus, being either a male or female offspring does not guarantee a diagnosis of schizophrenia when pathology supervenes, for one is just as apt to turn out a neurotic, chronic alcoholic, or to have an affective psychosis.

Again, the order of birth within a family, while it may predispose to stress of the kind postulated by Adler, seems to have nothing to do with the form of pathology which develops insofar as a mental hospital population is concerned. Potentially, a child born in any order is as much a candidate for schizophrenia as he is for an affective psychosis, etc. Thus, our findings in essence confirm that of Grosz and Miller, and Wahl.

It would appear to us at this time that in studying family dynamics in relationship to the causation of schizophrenia more attention should be given to the interpersonal nuances of family membership than to its structural aspects as such. Thus, it is conceivable that two families with identical structures can produce varying pathological products only one of which is schizophrenia.

REFERENCES

1. ADLER, A. What Life Should Mean To You. Boston: Little, 1931.
2. BOSSARD, J., & BOLL, E. S. Adjustment of siblings in large families. *Am. J. Psychiat.*, 1956, **112**, 889-892.
3. BOWEN, M. A Family Concept of Schizophrenia. *In* Jackson, D. D. (*Ed.*), *The Etiology of Schizophrenia.* New York: Basic Books, 1960.
4. GROSZ, H. J., & MILLER, I. Sibling patterns in schizophrenia. *Science*, 1958, **128**, 11-13.
5. HILGARD, J. R. Personal communication, 1962.
6. HILGARD, J. R., & NEWMAN, M. F. Evidence for functional genesis in mental illness: schizophrenia, depressive psychosis and psychoneuroses. *J. Nerv. Ment. Dis.*, 1961, **132**, 3-16.
7. JACKSON, D. D. Personal communication, 1962.
8. JONES, H. E. Order of Birth in Relation to the Development of the Child. *In* Murchison, C. (*Ed.*), *Handbook of Child Psychology.* Worcester, Mass: Clark Univ. Press, 1931. Pp. 204-241.
9. LIDZ, T. Schizophrenia and the family. *Psychiatry*, 1958, **21**, 21-27.
10. LIDZ, T., & FLECK, S. Schizophrenia, Human Integration, and the Role of the Family. *In* Jackson, D. D. (*Ed.*), *The Etiology of Schizophrenia.*
11. LIDZ, R. & LIDZ, T. The family environment of schizophrenic patients. *Am. J. of Psychiat.*, 1949, **106**, 332-345.
12. POLLOCK, H. M., MALZBERG, B., & FULLER, R. G. Hereditary and environmental factors in the causation of manic-depressive psychoses and dementia praecox. *Amer. J. Psychiat.*, 1940, **96**, 1227-1244.
13. WAHL, C. W. Some antecedent factors in the family histories of 392 schizophrenics. *Am. J. Psychiat.*, 1954, **110**, 668-676.

Agnews State Hospital
San Jose, Calif.

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 337-344.

THE EFFECT OF MEDIATORS IN CONCEPT IDENTIFICATION*

Department of Psychology, Washington University

DANIEL S. LORDAHL[1]

A. INTRODUCTION

In concept identification tasks with two dimensions relevant, as exemplified by Archer, Bourne, and Brown (1), Brown and Archer (4), and Bourne (2), a single relevant cue may be associated with two of four possible responses, the one correct response on a given trial being dependent upon the presence of an additional relevant cue. For example, if the stimuli are variations of geometric forms, color and shape of the forms might be selected as relevant dimensions. In this case, one of two responses (R_1 or R_2) out of four possible responses would always be correct for a red stimulus. Which response is correct would depend upon whether the stimulus is a red square (R_1 correct) or a red circle (R_2 correct). The remaining two responses would be appropriate for blue squares (R_3) and blue circles (R_4).

Bourne and Restle (3) applied a discrimination learning model to the concept task and treated the above complexity by assuming that subjects are required to solve two simple and separate discrimination problems before attaining a solution. This assumption seems applicable only where stimulus aspects are dimensionalized, i.e., when subjects have learned that certain cues are intimately related. For instance, again consider variations in the color of geometric forms as relevant cues to which subjects must learn to respond. In this case, two different cues such as red and blue are already related to one another, probably by a mediating response like the word "color".

The present study is concerned with the effect on subject's behavior of this organization of cues into dimensions of stimulus variation. The major hypothesis to be tested is: (*a*) subjects will respond with equal dependency upon two relevant cues which have been previously associated but will show unequal dependency upon two relevant cues not previously associated. The rationale for this hypothesis is that the learned relevance of a single cue

---

* Received in the Editorial Office on November 29, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The author wishes to acknowledge the graduate school of Washington University for their financial aid in obtaining equipment for this study.

(red for instance) is readily generalizable, via mediation mechanisms, to the associated cue (blue). Such generalization of relevance would be less when cues were not associated prior to the concept task.

A second hypothesis, based on the assumption that association of cues into sets (dimensions) can itself be a discrimination problem, would be: (*b*) subjects will have greater difficulty in a concept task when no prior dimensionalization of cues has been established than when such associations have been learned.

B. Method

1. *Subjects*

Forty-five subjects volunteered for the study. All were second semester introductory psychology students and received extra credit points for their service. Ten subjects were dropped for failure to solve the practice problem during the first session. One subject did not return for the second session.

2. *Materials*

All stimulus materials were photographed and projected onto a 6 × 8 inch translucent paper screen, about 3 feet in front of the subject, with a 16 mm. strip-film projector.

The stimuli for Problem 1 (a practice problem) were eight words printed in capital letters and constituted four bi-leveled dimensions: form, "CIRCLE" or "SQUARE"; texture, "ROUGH" or "SMOOTH"; color, "RED" or "BLUE"; and size, "LARGE" or "SMALL." For this problem, a strip-film of 128 frames was constructed, each frame containing four of the above words, one from each dimension, randomly arranged in a single printed line. All words appeared an equal number of times and redundancy was held at zero.

The strip-film for Problem 2 was constructed in a manner completely analogous with the above method, except that 8 symbols were substituted for the words used in the first problem. These eight symbols, shown in Figure 1, were randomly drawn and paired from a set of 30 symbols previously constructed by the experimenter.

Two paired associate lists were also filmed. These lists were identical with respect to the stimuli, both having four orders of presentation of the eight symbols used in Problem 2. As may be seen in Figure 1, List C contained a different nonsense syllable response for each symbol, while List E contained four nonsense syllables, each paired with two of the symbols.

For each of the two problem solving tasks, subjects pulled one of four levers mounted upright on the table before them on each trial. Feedback lights above each lever indicated the correct lever as soon as the subjects responded and remained on for 4 seconds at which time the next stimulus was projected on the screen. All responses were recorded by the experimenter and by an Esterline-Angus operations recorder. Partitions separated subject from experimenter and recording apparatus.

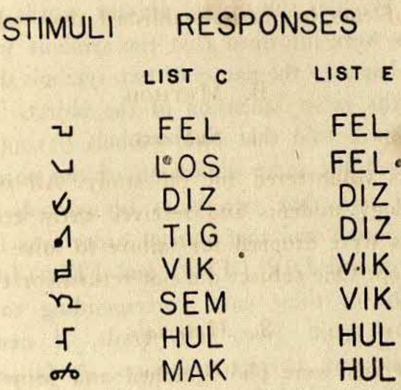


FIGURE 1
THE SYMBOLS USED IN PROBLEM 2 WITH THEIR RESPONSE NAME AS LEARNED ON TWO PAIRED ASSOCIATE LISTS

3. *Procedure*

The 34 subjects who completed the experiment had been randomly assigned to two groups, experimental (16 subjects) and control (18 subjects), and served on two consecutive days at about the same time each day. Pilot subjects had encountered great difficulty solving Problem 2 alone. For that reason, Problem 1 was used as a practice problem for both groups, the dimensions of form and texture being relevant. All subjects were instructed about the nature of the concept task and were told that two dimensions were relevant and two others irrelevant to solution. They were also told that the word order would vary from trial to trial but was definitely not important for solution. Response rate was self-paced and all subjects continued until they reached a criterion of 16 consecutive correct responses. For subjects who did not reach criterion within 128 trials, the Problem 1 film was repeated following a 60-second rest interval. The subjects were asked for a verbalization of the solution following their criterion run.

About three minutes after completion of Problem 1, all subjects received 16 trials on the paired associate lists, the experimental group practicing on

List E and the control group on List C. Items were presented at a 2-2-second rate with a six-second intertrial interval. Subjects were instructed to pronounce the nonsense syllable names.

At the start of the second session, all subjects were given a recall trial and seven relearning trials at the same rates on the appropriate list. All subjects were then told that they were to solve the same type of problem as on Day I, with the symbols substituted for the words used on Problem 1. The control subjects were informed that the symbols were paired but that similarities between shapes or the names of two symbols did not need to mean that they described the same dimension of the object. The subjects in the experimental group were told that two symbols having the same nonsense syllable name described the same dimension of the object.

All subjects were told that two pairs of symbols (dimensions) were again relevant for solution and that symbol order was irrelevant. For both groups, the symbols named DIZ (TIG) and FEL (LOS) were relevant, the four combinations of these pairs corresponding to the four response levers. Subjects were again given 256 trials, if needed, to reach the criterion of 16 consecutive correct responses. A 60-second rest followed Trial 128. Subjects were asked for verbal solutions following the criterion run.

C. Results and Discussion

Three subjects in the control group and one experimental subject failed to reach criterion for Problem 2, nor did these subjects verbalize an adequate solution to the problem at the end of 256 trials. The data from these subjects have been dropped from the following results.

Trials to criterion and errors to criterion were tabulated for each subject for each concept problem. On Problem 1, the mean trials to criterion (excluding the criterion run) were 72.33 for the experimental subjects and 72.40 for the controls. Mean errors to criterion were 41.53 for experimental subjects and 41.07 for controls.

On the paired associate learning tasks, all subjects reached a rather high level of proficiency, even the slowest subject attaining a level of three consecutive perfect trials. The mean errors per subject on Day I and Day II were, for the experimental subjects, 24.20 and 1.40, respectively. The corresponding means for the control group were 18.47 and .87. The Day I differences were not significant ($F = 1.51$, 1/28 *df*). The differences for Day II were not tested for significance.

The rather high degree of overlearning on the symbol names can be considered to minimize the effect of differential learning rates for Lists E and C while maintaining equal exposure to the symbols. Any bias introduced by learning the different lists can be expected to aid the control subjects on Problem 2. The mean errors in learning the two lists suggests this conclusion as well as the fact that control subjects are forced to discriminate among all eight symbols while the experimental subjects need not take careful note of differences between paired symbols in this learning task.

On Problem 2, the experimental subjects averaged significantly fewer trials to reach the criterion than did the controls. Table 1 shows a summary of the performances for the two groups, in terms of trials, errors, and correct responses to criterion. It can be seen that the experimental group had significantly fewer errors but also significantly fewer correct responses than the control group. The comparisons involving trials to criterion and correct responses to criterion both involved heterogeneity of variance. A logarithmic transformation of these data was used in comparing groups means. The transformation adequately reduced the heterogeneity, and variance tests then yielded F ratios of 6.07 for trials to criterion measures and 6.27 for the number of correct responses to criterion, both based on 1/28 *df*. Those F ratios are significant at the .025 level and confirm the reliability of the differences shown in Table 1.

TABLE 1
COMPARISON OF PROBLEM 2 PERFORMANCE

	Experimental ($N = 15$)		Control ($N = 15$)		
To criterion	Mean	*SD*	Mean	*SD*	*F*
Trials	71.27	38.35	118.20	61.57	5.86**
Errors	38.40	21.22	58.33	29.56	4.20*
Number correct	32.87	19.68	59.87	40.84	4.97*

* $p < .05$
** $p < .025$

In order to test hypothesis (*a*), each response for a given subject prior to his criterion run was placed into one of 16 categories. These categories were generated by combining the four types of stimuli (with regard to the correct response) with the four possible responses by the subject. The percentage of total responses for each group which fell into each category is presented in Table 2 (it should be noted that the discrepancies between the two sets of means are relatively small).

Certain of the cells in the matrices in Table 2 indicate the subjects' relative dependence upon one or the other of the two relevant symbol dimensions. Specifically, Cells 1-4, 2-3, 3-2, and 4-1 (TIG-DIZ errors) when compared with Cells 1-3, 2-4, 3-1 and 4-2 (LOS-FEL errors), show the degree to which subjects used the stimulus dimension TIG-DIZ over the dimension LOS-FEL. The mean number of responses from the former four cells is 23.50, summing over both groups while the mean for the latter cells is 13.90. This difference is significant beyond the .001 level, $F = 19.54$, with 1/29 *df*. This difference merely indicates that the subjects in both groups tended to find the relevance of one or both of the symbols TIG-DIZ prior to finding the relevance of the LOS-FEL symbols.

TABLE 2
PER CENT OF TOTAL RESPONSES FOR PROBLEM 2 CLASSIFIED BY STIMULUS TYPE AND RESPONSE LEVER

Correct response	Relevant symbols	Subject's response				Totals
		1	2	3	4	
		Experimental group				
1	LOS TIG	11.60	2.71	3.18	8.70	26.19
2	FEL DIZ	2.90	11.32	6.08	3.74	24.04
3	LOS DIZ	4.49	6.27	10.48	3.46	24.70
4	FEL TIG	5.80	3.18	3.37	12.72	25.07
	Totals	24.79	23.48	23.11	28.62	1069 responses
		Control group				
1	LOS TIG	11.62	2.88	2.82	7.73	25.04
2	FEL DIZ	2.42	13.37	4.91	4.34	25.04
3	LOS DIZ	3.33	5.81	12.75	3.21	25.10
4	FEL TIG	5.13	4.23	2.54	12.92	24.82
	Totals	22.50	26.28	23.01	28.20	1773 responses

Hypothesis (*a*) implies that experimental subjects would respond (prior to solution) to one symbol dimension, i.e., with equal weight given to each member of a relevant symbol pair, while control subjects would tend to cue their responses from only one symbol of a relevant pair. In order to consider only the first learned relevant cue, each subject was classified as a TIG-DIZ or a LOS-FEL responder on the basis of the numbers of errors of each type (as defined above). Eleven of the 15 subjects in each group fell into the TIG-DIZ class. For these TIG-DIZ subjects, the number of correct responses on Levers $1 + 4$ were compared with the number of correct responses on Levers $2 + 3$. If TIG and DIZ are given equal weight by the subjects, these two totals should not differ significantly (allowing for simulus frequencies). Expected values for these two totals were computed for each subject, adjusting for stimulus frequencies, and a χ^2 (corrected for discontinuity) was obtained for each subject. For the LOS-FEL responders, the same procedure was followed but the correct response totals compared were from Levers $1 + 3$ and $2 + 4$. The resultant χ^2 totals were 6.32 for the experimental group and 32.64 for the control group, each based on 15 *df*. The control group total is significant, $p < .01$, while the experimental group total is well within reasonable chance expectancy. The ratio of these totals is 5.16, which, when interpreted as an F ratio, is significant beyond the .01 level.

These data give clear support for Hypothesis (*a*). The experimental subjects show close conformity to the predicted equal weighting of dimensionalized relevant cues. The control subjects depart significantly from the equal weight hypothesis; moreover, they show significantly greater deviations from equal weight predictions than do the experimental subjects.

A reasonable conclusion appears to be that control subjects were learning to respond to the individual symbols and their combinations rather than to symbol pairs (dimensions) to a greater degree than were experimental subjects.

The ease of concept identification, with and without dimensionalized cues, as defined by the two conditions in the present study, would seem to derive from the notion that subjects in the concept task can learn about the relevance of a pair of cues from noting the relevance of one member of the pair. Note, however, that Hypothesis (*b*) cannot be unambiguously tested from the present data. Clearly, control subjects had more difficulty on the test task but the reasons for this difficulty might not derive *directly* from the lack of dimensionalization of stimuli. For example, control subjects may have spent time attempting to find paired symbols on the test task prior to "searching" for relevant cues.

D. SUMMARY

Two groups were used to study the effect on concept identification of dimensionalized versus undimensionalized stimuli. Both groups were given a practice problem on a concept task in which four bi-leveled stimulus dimensions were represented by eight words describing the characteristics of common objects. In the test task, eight symbols were used in place of the words. The experimental group first learned a common nonsense syllable response to each pair of symbols, a pair representing a stimulus dimension. The control group learned different nonsense syllables responses for each of the eight symbols.

The analysis of response patterns gave evidence which suggested that the control subjects were solving the problem in terms of the individual symbols while the experimental subjects were utilizing symbol dimensions.

The experimental group showed significantly fewer trials to reach criterion on the test task than the controls.

REFERENCES

1. ARCHER, E. J., BOURNE, L. E., JR., & BROWN, F. G. Concept identification as a function of irrelevant information and instructions. *J. Exp. Psychol.*, 1955, **49**, 153-164.
2. BOURNE, L. E., JR. Effects of delay of information feedback and task complexity of the identification of concepts. *J. Exp. Psychol.*, 1957, **54**, 201-207.
3. BOURNE, L. E., JR., & RESTLE, F. Mathematical theory of concept identification. *Psychol. Rev.*, 1959, **66**, 278-296.
4. BROWN, F. G., & ARCHER, E. J. Concept identification as a function of task complexity and distribution of practice. *J. Exp. Psychol.*, 1956, **52**, 316-321.

Department of Psychology
Washington University
St. Louis 30, Missouri

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 345-355.

PAIN SENSITIVITY OF AMPUTATION EXTREMITIES*[1]

New York University

HYMAN KORIN, SAMUEL A. WEISS, AND SIDNEY FISHMAN

A. INTRODUCTION

The systematic study of pain thresholds after amputation has lacked attention despite evidence that other cutaneous thresholds in amputation extremities are lowered. The altered cutaneous sensitivity has been reported in the early basic studies of Katz (3, 4) and more recently in the carefully controlled studies of Teuber (5) and Haber (2). These investigators have compared corresponding portions of the stump and intact limb in unilateral amputees. They have found that though the dense peripheral nerve supply of the missing extremity was lost, the stump showed a proximal-distal gradient of increased tactual sensitivity; the more distal the area the lower was the threshold. Teuber and Haber have concluded that the gradient may be attributed to a central reorganization by means of which the stump takes over the functional characteristics of the missing distal extremity.

After amputation of a limb, pain is invariably felt during the early post-surgical period. This pain usually subsides unless infection of bone or tissue, trauma to nerve ends, a painful neuroma, nerves caught in scar tissue, or psychogenic conditions occur to prolong it. When the stump is well-healed, however, occasional paresthesiae may occur. Some amputees indicate that they are aware of an increased sensitivity, when asked about it, which is felt as a tingling, numbness, tightness or even pain. These sensations are usually experienced intermittently in circumstances varying from emotional overexcitement to a change in the weather, and are often felt in a phantom limb as well as in the stump. Palpation of scar tissue may also sometimes

---

* Received in the Editorial Office on November 30, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This investigation was supported in part by a research Special Project grant from the Office of Vocational Rehabilitation, Department of Health, Education and Welfare. Portions of the study were presented at the convention of the Eastern Psychological Association, April, 1962. The authors are especially grateful to the following vocational counselors for aiding in the recruitment of amputees; Mr. Harry Katz, Mrs. Shirley Edelstein, Mr. Meyer Kaminsky, Mr. Israel Weissfeld, and Dr. Daniel Sanford of the New York City offices of the State Division of Vocational Rehabilitation; Mr. Charles Sadbury, Coney Island Hospital, Brooklyn, New York, and Mr. Cornelius Stewart, City Hospital at Elmhurst, Elmhurst, New York.

induce a slight tingling sensation. However, in most instances, there is no obviously greater sensitivity to pain in the stump than in the intact limb.

In the present study, patterns of pain threshold sensitivity in amputation lower extremities have been determined and these patterns have been related to various amputation conditions to ascertain whether support is given to a conception of either a peripheral or central origin or both. An additional purpose of the study was to secure information concerning tolerances to pressures, which might be useful in the design of prosthetic sockets for lower-extremity amputees. Although it is generally appreciated that various aspects of the stump possess a differential ability to tolerate socket pressures, no systematic exploration of this pattern has been previously undertaken. It was, therefore, considered likely that by "mapping" the pressure sensitivity of the stump, clues leading towards more rational shaping of sockets may be found.

B. METHOD

1. *Subjects*

The subjects were 44 recently amputated lower-extremity unilateral amputees. Twenty-four had had a below-knee amputation and twenty an above-knee amputation. All were examined just before being fitted with a prosthetic device. Each amputee had a well-healed stump. None had ever used a prosthetic device before. The age means of the two amputation groups were approximately 48 years. Although the age ranges were from 17 to 73, about 20 per cent of each group exceeded 60. Of the 24 below-knee subjects, 14 were diabetics, three had incurred a traumatic injury, two osteomyelitis, three a peripheral vascular disease, one a tumor, and one frostbite. In the above-knee group, 10 of the 20 subjects had a peripheral vascular disease, three had incurred a traumatic injury and two osteomyelitis. Fourteen subjects had a left limb amputated. Sixteen subjects were females.

2. *Apparatus*

Pain may be induced by the application of thermal, chemical, electrical or mechanical stimulation. A mechanical skin deformation type stimulation, of relatively deep tissue, was selected in the present study because it is most like the stimulation that an amputee experiences when he wears an artificial limb. Thus the instrument used to measure the pain threshold consisted of a piston-rod enclosed in a metal tube in which was mounted a precision spring. A plastic fitting, one-quarter inch square with bevelled edges, was attached to the tip of the rod. Upon pushing the fitting against a skin area, the force in pounds exerted was read from a scale on the metal tube which indicated the degree of compression of the steel spring. This scale registered from zero to 40 pounds in one-half pound units, by means of a fixed marker.

3. *Location of Testing Sites*

The testing sites were located on homologous distal and proximal aspects of the stump and intact limb (Figure 1). The distal sites were determined by drawing circles to form one-half inch bands, starting one inch from the distal end of the stump. The proximal bands were one-half inch bands drawn one inch from the perineum in the above-knee amputee and one inch from the medial tibial plateau in the below-knee amputee. The bands of the intact limb were drawn at the same height levels as those of the stump. All bands were then subdivided into one-half inch squares to accommodate the rod fitting. In each band, every odd numbered square served as a testing site. A counter-balanced design was used in selecting the sequence of testing the stump and the intact limb, the distal and proximal bands and the first site in each band. After a reading was taken on the first site of a band, the testing proceeded in a clockwise direction on the stump and in a counter-clockwise direction on the intact limb, so that the homologous areas of each extremity were examined in the same sequence.

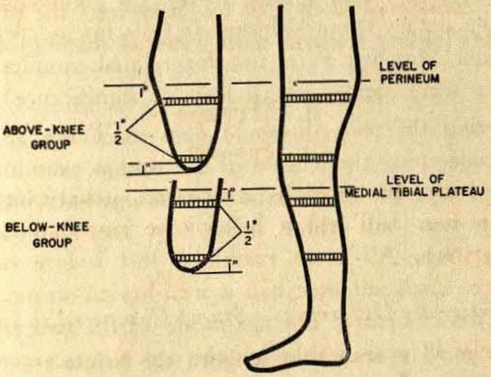


FIGURE 1
DISTAL AND PROXIMAL CIRCUMFERENTIAL BANDS ON STUMP AND HOMOLOGOUS AREAS OF INTACT LIMB IN ABOVE-KNEE AND BELOW-KNEE GROUPS

4. *Procedure*

To obtain the pain threshold measures, the piston-rod fitting of the gauge instrument was placed flat against a skin area by holding the instrument perpendicular to the testing site. The force was then increased by the examiner slowly and uniformly until the subject reported just when the area began to hurt. The subject was tested while he stood upright. He placed his limb against a rigid surface and changed his position as required until all the bands were tested.

The pain threshold scores of the distal and proximal bands were determined by calculating the mean of all measures of each band.

C. RESULTS

The distal and proximal band scores of both the stump and the intact limb, in the below-knee and above-knee groups, were all skewed to the right. Nonparametric tests were therefore applied to test the significance of differences between the various data (Mann-Whitney *U* tests for independent samples of data; the Wilcoxon Matched Pairs test for related samples of data. Two-tailed probabilities were applied in all tests of significance).

Before considering the pain threshold data in Table 1, several findings should be noted concerning the lengths of the stumps examined. The distance between the distal and proximal bands (each one-half inch wide) in the below-knee group was only three inches; in the above-knee group this distance was six inches.[2]

1. *Inter-Extremity Differences—Stump Compared to Intact Limb*

In Table 1, it may be seen that in both the below-knee and above-knee subjects the distribution of threshold scores on the stump and intact limb differed markedly whether the distal or proximal bands were compared. On the distal band of the stump, for example, the median score of the below-

TABLE 1
MEDIAN AND RANGE OF THE PAIN THRESHOLD SCORES ON THE STUMP AND INTACT LIMB OF 24 BELOW-KNEE AND 20 ABOVE-KNEE AMPUTEES

		Stump		Intact limb	
		Distal	Proximal	Distal	Proximal
Below knee	Median	7.53	12.04	17.88	17.33
	Range	3.50–40.00	7.63–40.00	9.33–40.00	9.33–40.00
Above knee	Median	5.70	9.13	14.29	14.29
	Range	1.71–14.60	3.50–27.00	6.33–28.35	7.35–33.67

---

[2] Below-knee stumps are usually amputated five to six inches in length by surgeons; above-knee stumps are amputated ten to twelve inches in length. Stumps of these lengths tend to develop the least circulatory problems and from a prosthetic point of view give satisfactory leverage.

knee and above-knee subjects was 7.53 and 5.70 pounds. By contrast, the median scores of the thresholds in the distal band of the intact limbs of the two amputation groups were about 10 pounds greater than the median scores on the stump, being 17.88 pounds in the below-knee group and 14.29 pounds in the above-knee group.

Turning from the distal to the proximal bands, in the two amputation groups, it is also evident that the thresholds tended to be lower on the stump than on the intact limb. The difference between the proximal stump and proximal intact limb median scores, however, was only about five pounds instead of ten pounds as in the distal band comparisons.

The scores of the individual subjects are not given in the table, but it is to be noted that of the total of 24 below-knee and 20 above-knee subjects only two of the below-knee group had a higher threshold on the distal band of the stump than on the distal band of the intact limb; and only one of the latter two subjects as well as another below-knee subject had a higher proximal band threshold on the stump than on the intact limb. Thus, all the above differences in the two amputation groups between stump and intact limb, distal or proximal, were significant at better than the .001 level.

2. *Intra-Limb Differences—Distal Compared to Proximal Bands*

The second set of differences concern those between the distal and proximal bands, first on the stump and second on the intact limb. On the stump, in both amputation groups, the distal band threshold scores were significantly below those of the proximal band ($p < .001$). On the intact limb, the differences between the scores of the distal and proximal bands lacked significance. In an additional comparison, the above-knee subjects tended to have lower thresholds than the below-knee subjects in the distal and proximal bands, both on the stump, and intact limb; the differences between the two amputation groups, however, were significant only at the proximal end of the stump.

To summarize these findings, both the below-knee and above-knee groups had significantly lower pain thresholds at the distal and proximal ends of the stump than in comparable areas of the intact limb; on the stump, the distal thresholds were significantly lower than those at the proximal end.

3. *Pain Thresholds of Lateral, Anterior, Medial or Posterior Aspects of the Stump and Intact Limb*

The aspect of a circumferential band (lateral, anterior, medial or posterior) in which the thresholds were measured was recorded. The average score in each aspect was considered to be its threshold. To determine whether the thresholds of these various aspects differed at both ends of the stump, the following comparisons were made: first, intra-band (in each band of the stump and intact limb) between the threshold scores of the four aspects; second, intra-limb (stump or intact limb) between the corresponding aspects of the distal and proximal bands; third, inter-limb (stump against intact limb) between the thresholds of the corresponding distal aspects and also the corresponding proximal aspects.

The results of the below-knee and above-knee groups are presented separately in Table 2. Subdivisions indicate the data of the stump and intact limbs.

a. Below-knee group

(1). *Stump.* In both the distal and proximal bands, the distribution of the thresholds of pain sensitivity of the various aspects of the stump were not significantly different. All aspects of the distal band, however, had significantly lower score distributions ($p < .01$) than the corresponding segments of the proximal band.

(2). *Intact limb.* In both the distal and proximal bands, the medial aspect had a significantly lower score distribution than the lateral, anterior or posterior aspects ($p < .01$ to $p < .05$). The differences between the corresponding parts of the distal and proximal bands were not significant.

(3). *Stump-intact limb.* All the aspects of the stump, distally and proximally, had significantly lower scores than those of the corresponding aspects of the intact limb ($p < .01$ each comparison).

b. Above-knee group. The pattern of scores on the stump and intact limb in this group differed from that in the below-knee group.

(1). *Stump.* In the distal and proximal bands, instead of there being no significant differences between the thresholds of the various aspects of the stump as occurred in the below-knee group, the medial area had significantly lower scores than the lateral, anterior or poterior areas ($p < .01$ to $p < .02$). In the proximal band, furthermore, the distribution of scores of the anterior segment was significantly below that of the lateral segment ($p < .05$). In turn, both the anterior and lateral areas had significantly lower scores than the posterior area ($p < .01$ to $p < .02$). Despite the small differences between the medial scores of the corresponding distal and proximal aspects, the thresholds on the distal band tended to be below those of the proximal band. Thus, the differences between all the distal and proximal segment scores were also significant at better than the .01 criterion.

(2). *Intact limb.* In the distal band the medial and posterior scores were

TABLE 2

MEDIAN AND RANGE OF PAIN THRESHOLD SCORES ON LATERAL, MEDIAL, ANTERIOR AND POSTERIOR ASPECTS OF THE STUMP AND INTACT LIMB IN BELOW-KNEE AND ABOVE-KNEE GROUPS*

		Below-knee group				Above-knee group			
		Stump		Intact limb		Stump		Intact limb	
		Distal	Proximal	Distal	Proximal	Distal	Proximal	Distal	Proximal
Lateral	Median	7.25	10.13	18.75	18.75	7.25	9.50	15.84	15.00
	Range	3.75–40.00	6.00–40.00	8.50–40.00	9.75–40.00	1.00–14.00	4.25–22.00	7.00–33.00	7.00–34.50
Anterior	Median	7.63	11.67	17.25	18.25	6.00	7.00	14.25	13.75
	Range	3.50–40.00	4.67–40.00	8.00–40.00	6.00–40.00	2.00–31.00	4.00–31.00	6.33–28.35	7.17–33.25
Posterior	Median	6.00	10.38	15.25	15.00	4.50	6.00	12.00	10.67
	Range	2.00–40.00	4.00–40.00	6.75–40.00	7.00–40.00	1.50– 9.00	3.75–20.00	5.00–18.38	5.00–28.00
Medial	Median	7.50	13.50	17.25	18.75	5.50	14.25	11.00	15.50
	Range	2.33–40.00	7.00–40.00	8.67–40.00	10.00–40.00	1.00–15.50	2.75–34.50	6.00–35.00	7.67–40.00

* Significance of differences between distributions of scores of various aspects are given in the text.

significantly below the lateral or anterior scores ($p < .01$). In the proximal band, the medial aspect had thresholds that were significantly below those of the other aspects ($p < .01$). In a comparison of the corresponding aspects of the distal and proximal bands, only one difference was significant; the posterior aspect of the distal band had significantly lower scores than that of the proximal band ($p < .01$).

(3). *Stump and intact limb.* The medial, lateral and anterior aspects of the distal and proximal bands of the stump had significantly lower score distributions than the corresponding parts of the intact limb ($p < .01$), whereas the scores of the corresponding posterior parts of the proximal bands of the stump and intact limb were not significantly different.

4. *Relation Between Age, Laterality, Etiology, Stump Circumference and the Mean Threshold Score of the Distal or Proximal Band*

In additional analyses the following results were obtained: low correlations were found between the pain thresholds and age; no significant differences were found between the scores of the right and left limbs when the thresholds of either the stump or intact limb were compared; the males were significantly different from the females, only at the distal end of the below-knee stump. In the latter result the females tended to have the lower thresholds ($p = .05$).

Twenty subjects, 14 below-knee and six above-knee, had diabetes. These patients tended to have higher thresholds on the distal end of the stump than the 24 other subjects significant at the .02 level. In the below-knee group, nine of the 14 diabetic subjects (64 per cent) were above the median (7.53) of the group against only 3 of the 10 non-diabetics (33 per cent). In the above-knee group, five of six diabetics (83 per cent) were above the median (5.75) against only five of 14 non-diabetics (36 per cent). In the non-diabetic group, 13 peripheral vascular disease subjects tended to have low threshold scores; the scores of four of these who had arteriosclerosis, however, were relatively high. The subjects whose amputations were brought about by other etiologies, consisting of a total of 11, tended to have low thresholds.

The pain threshold scores were also correlated with measures of the circumference of the stump taken at the distal and proximal ends because it was observed that a few heavy tissued subjects had high thresholds. In the above-knee group, a rank difference correlation of .42 between the circumference and pain scores at the distal end was significant ($p < .05$). In the below-knee group, the correlation at the distal end was only .29. When, however, one rather emaciated diabetic subject who had a high pain threshold was omitted from the sample, the correlation increased to a significant .43 ($p < .05$).

Seven above-knee and three below-knee subjects who had a mildly sensitive scar to palpation tended to have low thresholds both on the stump and on the intact limb. Although six of these subjects had either adherent or invaginated scar tissue, five other subjects who had such tissue had high thresholds. Thus, the differences between the adherent-invaginated and non-adherent-non-vaginated groups were not significant.

D. Discussion

The results of the present study follow a pattern that has been reported in studies of tactile sensitivity of the amputation extremity by other investigators. The stump is more sensitive than the intact limb; the distal end of the stump is markedly more sensitive than the proximal end. In the tactile studies this altered sensitivity has been explained by a "central" theory. The rationale of that conception takes into account that the hand or foot of a normal limb has greater nerve supply, greater cerebral representation and greater tactile sensitivity than the more proximal portions of the limb. The heightened tactile sensitivity of the distal normal extremity, therefore, may be due to either the eccentricities of peripheral or central nervous organization or to both. When a limb is amputated, the distal sensory end organs are lost. Presumably, the heightened proximal-distal gradient of tactual sensitivity of the stump may be attributable to a reorganization of the central patterns of activity. By means of this reorganization, furthermore, it is as if the functional characteristics of the missing hand or foot have been referred to the stump. The adherents of the theory hold that the greater sensitivity of the stump than the intact limb, even at the proximal end, distant from the amputation, is evidence against the possibility that the sensitivity occurs because the nerve supply has been made more dense by post-operative distortions of the skin.

What theory best accounts for the heightened pain sensitivity? Although pain is poorly developed in its cognitive qualities, a central organic reorganization theory similar to that discussed above might also explain the heightened effect in this modality. Pain is indeed mediated by central processes, is probably localized in the body schema by related tactile and pressure impressions (1), and when occurring in a phantom limb has been alleviated at least transiently by the excision of appropriate cortical tissue of the post-central gyrus. A psychogenically oriented theorist might suggest, however, that the heightened pain sensitivity is induced by the mental state of the subject. The experience of pain, investigators have found, has two components (6). There is perception of it and the reaction to this perception depending on the subject's attitude. Thus, the stump may be cathected emotionally, so that the amputee says "hurts" more readily when force is applied to that area rather than elsewhere on the body. Several findings in the present study, however, indicate that such an attitude is not of importance in most of the subjects and that instead the condition of the stump is primary. This interpretation is suggested by the uniformity with which the thresholds were markedly lower at the distal end of the stump close to the surgery regardless of the personality of the subject; by the finding that some of the low-threshold subjects also had slight scar sensitivity, and by the significant correlation between the thresholds and the circumference of the stump. Furthermore, the finding that diabetic subjects had the least sensitive thresholds probably also occurred because of peripheral, rather than psychogenic or other central influences. These subjects are more likely to develop neuropathies than subjects who have had amputations because of other diseases or trauma. Studies have shown that even diabetics who lack a demonstrable neuropathy may have diminished nerve conduction.

If the peripheral environment of the stump may be a source of the heightened pain sensitivity, it is essential to examine the details of the amputation surgery. In the usual procedure, the nerves are dissected and permitted to retract to a position approximately one inch from the end of the stump. This retraction allows the formation of terminal end bulbs or neuroma to remain well-padded by surrounding muscle without becoming caught in scar tissue. To the degree that the nerve fibers lack padding despite all precautions, the subject may be more sensitive to pain at the distal end of the stump, or even at the proximal end when the stump is short. But how can the lowered threshold at the far-distant proximal aspect of a long stump be explained by a peripheral process? It is known that the structures of the stump alter because the longitudinal stress applied at the proximal end by the muscles is greater than at the distal end. The muscles and tendons of the stump atrophy and its bones nearly always become less calcified than those of the intact limb. The lack of calcification and atrophy which usually is more prominent at the distal end than at the proximal end, curiously mirrors that of the gradient of heightened tactile and pain sensitivity. This pattern of altered structure suggests that a "stress and strain" rearrangement on the stump may distort the tissues and its nerve supply and may furthermore produce subtle physiochemical changes which lead to neuronal hyperexcitability, so that the proximal-distal gradient of heightened sensitivity to both pain and tactile stimulation occur. The pain threshold may not always be lowered by such a distortion. The relatively high threshold on the posterior segment of the proximal band in the above-knee group appears to result from a tendency of the hamstring muscle to atrophy, become flabby and by bunching to give the posterior area a supply of well-padded tissue.

With reference to the second goal of the research, the matter of artificial limb socket design, the data confirmed the clinical impression of a significant reduction in stump sensitivity as one proceeds proximally along the stump. The present prosthetic practice of designing for "weight-bearing" in the proximal areas of the stump is clearly supported. The proximal tissues of the below-knee stump apparently tolerate pressure equally well on all four aspects (posterior, lateral, anterior and medial), whereas the posterior aspect of the above-knee stump is considerably less sensitive than the other three areas. These findings have clear implications for socket design and it may be that further analysis of the patterns of pain thresholds may lead to more specific suggestions for the arrangement of reliefs and build-ups in prosthetic sockets.

E. SUMMARY

The findings of the present study indicate that the pain sensitivity of the amputation limb is heightened; the distal end was more sensitive than the proximal end. Although a central reorganization may explain this pattern of sensitivity, peripheral factors play a critical role. Implications for the design of prosthetic devices have also been noted.

REFERENCES

1. FULTON, J. F. Physiology of the Nervous System. 3rd ed. New York: Oxford, 1949.
2. HABER, W. B. Effects of loss of limb on sensory functions. *J. of Psychol.*, 1955, **40**, 115-123.
3. KATZ, D. Psychologische versuche mit amputierten. *Z. Psychol.*, 1920, **85**, 83-117.
4. ———. Zur psychologie des amputierten und seiner prothese. Beihefte, *Z. angew Psychol.*, 1921, **25**, 6-118.
5. TEUBER, H. S., KRIEGER, H. P., & BENDER, M. S. Reorganization of sensory function in amputation stumps: Two-point discrimination. *Fed. Proc.*, 1949, **8**, 156.
6. WOLFF, H. G., & GOODELL, H. The relation of attitude and suggestion to the perception of and reaction to pain. A. Research. *Nerv. Ment. Dis., Proc.*, 1943, **23**, 434.

Amputee Psychology Research Project
New York University
252 Seventh Avenue
New York 1, New York

Published in abstract and in The Journal of Psychology, 1963, 55, 173-182.

THE EFFECTS OF SMALL AND LARGE GROUP INSTRUCTION ON LEARNING OF SUBJECT MATTER, ATTITUDES, AND INTERESTS*

Department of Psychology, Wisconsin State College

JOHN J. [illegible]

A. PROBLEM

Since the end of World War II, there has been increasing concern [illegible] for the problem of providing [illegible] at all levels [illegible] has been [illegible] the increasing enrollments. The question of the effectiveness of instruction in a larger or smaller class [illegible] and technological innovations which have facilitated instruction of greater numbers of learners. Teaching machines and television are examples of the latter. [illegible] attention has been directed [illegible] outcomes [illegible] Nachman and Opochinsky [illegible] psychology [illegible] concluded that students learned the cognitive content of college courses as well in large as in small sections; that students prefer instruction in small classes, and that the instructor may be a more important factor than the size of the instructional group. [illegible] found that three instructional methods in a child development course, lecture-centered, case-centered, and group-centered, yielded no significant differences in learning of subject matter and attitudes toward children. [illegible] reported that a lecture method of teaching to large groups in a speech course coupled with individual meetings of small sections for applied speech activities resulted in equal mastery of speech principles and skills [illegible] small classes taught in the traditional manner.

The present study was concerned with learning outcomes in small and large class sections of educational psychology.

---

* Received in the Editorial Office on December 4, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The research reported herein was supported by a grant from the Board of Regents of Wisconsin State Colleges.

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 357-362.

THE EFFECTS OF SMALL AND LARGE GROUP INSTRUCTION ON LEARNING OF SUBJECT MATTER ATTITUDES, AND INTERESTS*

Department of Psychology, Wisconsin State College

JOHN F. FELDHUSEN

A. PURPOSE

Since the end of World War Two increasing attention has been turned to the problem of class size in relation to learning outcomes. At all educational levels and in most subjects the need to increase the pupil-teacher ratio has been forced upon schools by the shortage of teachers and by increasing enrollments. The question of the effectiveness of instruction in a larger or smaller class has also been explored as an outgrowth of technological innovations which have facilitated instruction to greater numbers of learners. Teaching machine and televised instruction are examples of the latter. Most often attention has been directed to cognitive learning outcomes, namely the mastery of the subject matter of the course.

Nachman and Opachinsky (5) found no difference in final examination performance between students enrolled in large and small classes of general psychology. Siegel, Macomber, and Adams (6) concluded that students learned the cognitive content of college courses as well in large as in small sections, that students prefer instruction in small classes, and that the instructor may be a more critical factor than the size of the instructional group. Leton (4) found that three instructional methods in a child development course, lecture, case-centered, and group-centered, yielded no significant differences in learning of subject matter and attitudes toward children. Giffin and Bowers (3) reported that a lecture method of teaching to large groups in a speech course when coupled with additional meetings in small sections for applied speech activities resulted in equal mastery of speech principles and skills with small classes taught in the traditional manner.

The present study[1] was concerned with learning outcomes in small and large class sections of educational psychology.

---

* Received in the Editorial Office on December 5, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The research reported herein was supported by a grant from the Board of Regents of Wisconsin State Colleges.

B. METHOD

1. *Subjects*

The subjects were college sophomores enrolled in a course in educational psychology. A total of 44 subjects, 32 women and 12 men, were used.

2. *Procedure*

During the first week of a semester the following tests were administered: (*a*) comprehensive test on concepts and principles of educational psychology consisting of 104 multiple-choice items; (*b*) the Otis Quick-Scoring Mental Ability Gamma Test Form Am; (*c*) the Minnesota Teacher Attitude Inventory; and (*d*) the California Vocational Interest Analysis Personal-Social which yields a measure of interest in teaching. A group of 22 students from a large class of 72 students and 22 from a small class of 28 were selected by matching on the basis of the Otis raw scores. There were 16 women and six men in each group. This selection of experimental groups involved no singling out of the subjects for individual or experimental group treatments. Each group of 22 students simply remained in his somewhat larger small or large class.

During three periods each week the large group of students was taught with lectures and demonstrations and a very small amount of discussion and attention to student questions. For a fourth meeting the class was divided into two smaller groups which met at different times but with the same teacher. This meeting was used for discussion related to lectures and readings, experimental demonstrations, and discussion of test results. The small group of 28 students also met four periods a week and was taught with the same lectures and demonstrations.

At the end of the semester the test on concepts and principles of educational psychology, the Minnesota Teacher Attitude Inventory, the California Vocational Interest Analysis Personal Social, and a teacher and course evaluation scale of 20 items, developed by the author, were administered to all students in both groups. In addition, a special scale consisting of five items was administered to all the students in the large group to assess their reactions to the large class.

C. RESULTS

Post-test scores on the 104-item test of knowledge of concepts and principles of educational psychology were used to test a null hypothesis which stated that there would be no significant difference in the amount learned between groups of students taught in large and small sections. In Table 1 it is seen that the mean for the 22 students in a large class was 75.32, for the 22 in a small class, 78.23. The *t*-test for matched groups (2) yielded a *t* of 1.65. This *t* is not significant at the .05 level, and hence the null hypothesis is not rejected. The amount learned by students in a large section is apparently not significantly different from the amount learned by students in a small section.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF PRETEST AND POST-TEST SCORES FOR STUDENTS TAUGHT IN SMALL AND LARGE GROUPS

	Students in large group (N = 22)		Students in small group (N = 22)	
Variable	Mean	SD	Mean	SD
Otis mental ability raw scores	63.77	7.29	62.95	7.78
Educational psychology pretest	66.86	3.69	68.59	5.19
Educational psychology post-test	75.32	5.25	78.23	7.15
Minnesota teacher attitude inventory pretest	77.05	24.72	72.64	30.84
Minnesota teacher attitude inventory post-test	97.41	23.16	102.09	27.98
Interest in teaching pretest	23.50	6.12	26.55	6.20
Interest in teaching post-test	24.91	8.15	27.77	7.79
Evaluation of course and teacher	76.64	8.20	77.18	10.40

The pretest and post-test scores on the Minnesota Teacher Attitude Inventory were used to test a null hypothesis which asserted that there would be no significant difference between terminal measures (end of course) of students' attitudes in large and small classes. The specific area of attitudes would, of course, relate to child development and to teaching methods. The MTAI pretest means were 77.05 for the large class, 72.64, for the small group. The post-test means were 97.41 and 102.09 in the same order. The difference between the pretest means, while not significant at the .05 level, were deemed sufficiently great to preclude an analysis of post-test MTAI scores without taking cognizance of the pretest differences. Furthermore, the MTAI pretest and post-test scores for the two groups combined revealed a correlation of .74. The difference in pretests plus the substantial correlation lead to the selection of an analysis of covariance design (2) to test the hypothesis. The difference between the adjusted MTAI post-test means of the large and small groups, was found to be significant at the .05 level according to the analysis of covariance as reported in Table 2.

TABLE 2
ANALYSIS OF COVARIANCE OF MINNESOTA TEACHER ATTITUDE INVENTORY SCORES FOR STUDENTS TAUGHT IN LARGE AND SMALL GROUPS WITH *A* PRETEST OF THE INVENTORY AS *A* COVARIATE

Source	Sum of squares of errors of estimate	Degrees of freedom	Mean square deviation	*F*
Total	14522	42		
Within	12532	41	306	
Adjusted means	1990	1	1990	6.50[1]

[1] Significant at .05 level

While the changes from pretest to post-test in both experimental groups for the test of knowledge of concepts and principles and for the MTAI were significant at the .05 level, the change in interest in teaching was not significant from the beginning to the end of the semester. Both groups showed a very small gain. The pretest means of 23.50 and 26.55 for the large and small classes were not significantly different from one another according to the *t*-test which was used. The post-test means, 24.91 and 27.77 in the same order, also did not differ significantly.

The evaluation of the course and teacher by the students produced quite similar means, 76.64 for the large group, 77.18 for the small group. The difference was again tested with the *t*-test and found not to be significant.

Finally, a five-item opinion scale was administered to all the students in the large group. The large group consisted of 72 students but only 65 were used in this analysis of opinions because seven papers were defective in various ways which were judged to be invalidating. This analysis is reported in Table 3. The questions and options are stated together with the percentage for each response. A composite analysis for the five questions showed 13.6 per cent of the students selecting the first option which favored the large class, 46.6 per cent selecting the middle or neutral category, and 39.8 per cent the third category which favored the small class.

TABLE 3

EVALUATIONS OF THE LARGE CLASS ARRANGEMENT BY SIXTY-FIVE STUDENTS

1.	How well did you do (grades) in this larger class?
3%	Better than I usually do in smaller classes
55%	About the same as usual
42%	Not as well as in smaller classes
2.	How do you feel your participation in class activities such as question period, discussion, or other group activities was affected by the overall arrangement?
5%	I got to participate more
51%	About the same as usual
44%	I did not get to participate as much
3.	How much do you feel you learned in this larger class?
17%	More than I usually do in smaller classes
58%	About the same as in smaller classes
25%	Less than in smaller classes
4.	How do you feel that your motivation to study or learn was affected by this larger class?
17%	Motivation was greater than in small classes
44%	Same as in smaller classes
39%	Not as great as in smaller classes
5.	If you had a choice of taking a course in a section of 35 or in a section of 70, which would you choose?
26%	Large class
25%	No preference
49%	Smaller class
6.	Composite analysis for the five questions listed above.
13.6%	Favorable to large class
46.6%	Middle or neutral
39.8%	Unfavorable to large class

D. DISCUSSION

The results of this investigation are consistent with numerous other studies in which cognitive learning or mastery of subject matter seems to proceed as well for students enrolled in a large class as for students in a small class. However, a more adequate delineation of desirable cognitive learning outcomes is needed. Such a framework as would be afforded by the work of Bloom and others (1) would provide a more adequate basis to specify desirable outcomes. This should precede any further effort to develop or validate the terminal measures of learning for a study of this type.

The California Vocational Interest Inventory Personal-Social has not been studied intensively to determine its validity or power to measure interest in teaching, but it does purport to do this. For whatever validity the scale may possess, the results of this study indicate that interest in teaching is not adversely affected by instruction in a small or large group.

The Minnesota Teacher Attitude Inventory has been studied intensively, and its relationship to various aspects of teaching performance has been described. The results of this study indicate that students enrolled in large or small groups make significant gains in the quality of their attitudes as measured by the MTAI. However, those who are enrolled in small classes in educational psychology have more favorable attitudes at the end of the course. Possibly the small class affords teachers a better opportunity to establish rapport with students and thereby to teach or inculcate attitudes which are consistent with those measured by the MTAI.

Finally, it was found that only a few students prefer the large class, a large number are neutral or indifferent, and a smaller number prefer the small class. The latter, however, constitute 40 per cent. This study is consistent with others in finding this substantial number of students who prefer the small class.

E. Conclusions

A group of 22 students enrolled in a large class of 72 students in educational psychology learned the subject matter as well as a matched group of 22 students enrolled in a smaller class of 28 students. Terminal measures of interest in teaching and evaluation of the course and teacher revealed no significant differences between the groups. However, terminal MTAI measures were significantly greater for the students enrolled in a small class. It was also noted that 40 per cent of all the students in the large class indicated a preference for a small class.

References

1. Bloom, B. S., *Ed.*, Taxonomy of Educational Objectives: Handbook I: Cognitive Domain. New York: Longmans, Green, 1956.
2. Edwards, A. L. Experimental Design in Psychological Research. New York: Holt, Rinehart, & Winston, 1960.
3. Giffin, K., & Bowers, J. W. An experimental study of the use of lectures to large groups of students in teaching the fundamentals of speech. *J. Educ. Res.*, 1962, **55**, 383-385.
4. Leton, D. A. An evaluation of course methods in teaching child development. *J. Educ. Res.*, 1961, **55**, 118-122.
5. Nachman, M., & Opachinsky, S. The effects of different teaching methods: A methodological study. *J. Educ. Psychol.*, 1958, **49**, 245-249.
6. Siegel, L., Macomber, F. G., & Adams, J. F. The effectiveness of large group instruction at the University level. *Harvard Educ. Rev.*, 1959, **29**, 216-226.

Department of Education
Purdue University
Lafayette, Indiana

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 363-369.

MEASUREMENT OF PALMAR SWEATING: EFFECT OF REPEATED MEASUREMENT FROM THE SAME AREA*

George Peabody College and Southern Illinois University

H. Carl Haywood and Donald J. Shoemaker

A. Introduction

Studies utilizing the colorimetric measure of palmar sweating as an index of "anxiety" or "autonomic arousal" have produced encouraging but often inconsistent results. At least one of the reasons for the inconsistency in results appears to be the methodological problems involved in the use of the colorimetric method. Many investigators (1, 3, 4, 5, 6, 7) have demonstrated an increase in palmar sweating associated with the imposition of some anxiety-arousing stimulus situation into the subject's life. In one such study (4) palmar sweating was measured before and after the imposition of a supposed arousal-producing situation (experimental group), and before and after a supposed nonarousing situation (control group). In this study, both the control group and the experimental group showed a decrease in palmar sweating, with, however, the experimental group showing significantly less decrease. The design of the study involved taking successive measures from the same finger. From the obtained decrease in both groups, the present writers hypothesized that the measurement technique itself in the colorimetric method leads to a decrease in measurable palmar sweating if used repeatedly over short time intervals on the same skin area. Two studies were designed to test the hypothesis and to provide information concerning the reliability of the colorimetric method of measuring palmar sweating.

B. Method and Procedure

In the first study successive palmar-sweat measurements were taken from resting subjects, using two fingers at a time, with one finger constant throughout the series, that is, repeated measurements from the same finger, and the other finger varied. This procedure made it possible to compare successive measurements taken from different fingers but taken simultaneously with the control

---

* Accepted for publication by J. McV. Hunt of the Editorial Board. Received in the Editorial Office on December 5, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

finger, thus reducing variation due to time differences. This would also control for adaptation effects.

The subjects were 18 University of Illinois undergraduates and graduates. They were equally divided between males and females.

Upon entering the experimental room, each subject was given a 20-minute adaptation period. During this time the experimenter engaged each subject in light conversation, avoiding obviously emotional topics and controversial issues, and explained the procedure. Four palmar-sweat measurements were then taken at 10-minute intervals. Between palmar-sweat measurements, the subjects rested. The first measurement was taken from fingers two and three; the second from fingers two and four; the third from fingers two and one; and the fourth from finger two and the thumb. The right hand was used in all cases. This procedure means that whatever central adaptation occurs is expected to occur equally in the constant (previously used) finger and in the previously unused fingers, since the adaptation period is the same. Any differences in the amount of sweating between the constant finger and the previously unused finger on a given occasion of measurement must, then, be due to the process of measurement itself, and not to adaptation phenomena. The procedure for taking palmar-sweat measurements was that procedure recommended by Mowrer (9). Densitometer readings were made within 24 hours following the experimental procedure.

In the second study, 20 University of Illinois undergraduates followed exactly the same procedure, with two major exceptions: (*a*) the order of measuring fingers was varied systematically in such a way that each of the five fingers was used as the constant finger with four subjects, and all orders of measuring five fingers, two at the time with one held constant, were used; (*b*) instead of using paper impregnated with tannic acid for taking the palmar-sweat blot, the blots were made on photographic film washed in sodium thyiosulphate and soaked in a five per cent tannic acid solution, a procedure recommended by Brutten (2).[1] The first modification was designed to control for any effect of possible geographic differences between the fingers in amount of sweating and to control for the possible interaction of order effect with such geographical differences.

---

[1] This makeshift procedure is no longer necessary. Commercially-prepared blotting film, ready for use, as well as an automatic fingerprinter and an improved densitometer, are now available from Lab-line Instruments, Inc. Chicago 22, Illinois. The use of these techniques further improves the reliability of palmar-sweat measurement.

C. RESULTS

The results of both studies are presented in Table 1. Higher densitometer readings indicate less sweating since the data are expressed in microamperes transmittance. Before each series of readings the intensity of the light in the densitometer used to measure the blots was such that an unblotted portion of the prepared paper or film gave a densitometer reading of 10.0 microamperes. It can be seen from Table 1 that the tendency for the blots taken from previously used (contant) fingers to become lighter with successive measurement is greater than the small adaptation effect revealed in the blots taken from previously unused fingers. In the first study the mean of all measurements taken from previously used fingers was 5.39 microamperes transmittance, while the mean of all measurements taken from previously unused fingers was 4.02. A subject-by-subject comparison of these data yields a t value of 7.04 ($p < .001$). In the second study the comparable means are 8.18 for all measurements taken from previously used fingers, and 7.10 for measurements taken from previously unused fingers. The resulting t value is 4.48, with 19 degrees of freedom ($p < .001$).

TABLE 1

DENSITOMETER READINGS OF PALMAR-SWEAT BLOTS OBTAINED THROUGH SUCCESSIVE USE OF THE SAME FINGER AND THROUGH SUCCESSIVE USE OF DIFFERENT (PREVIOUSLY UNUSED) FINGERS

(Measures for each occasion were obtained simultaneously—occasions were separated by 10-minute intervals.)

Occasion of measurement	Constant finger			Previously unused finger				
	Finger no.	Mean	*SD*	Finger no.	Mean	*SD*	Mean difference	t^{a}
				Data from Study I				
A	2	4.06	1.54	3	3.83	1.87	0.23	0.98
B	2	5.27	1.65	4	3.61	1.83	1.66	5.56**
C	2	5.61	1.70	1	4.25	1.75	1.36	3.62*
D	2	5.33	1.52	thumb	4.31	1.73	1.02	2.97*
				Data from Study II				
A		6.65	2.11		7.18	2.00	—0.53	(1.58)†
B		7.65	1.62		7.18	1.99	0.47	1.03
C		8.08	1.82		6.88	2.41	1.20	3.19*
D		8.80	1.39		7.62	1.55	1.18	4.99**

[a] Probabilities given are for one-tailed tests with 17 *df* (Study I) and 19 *df* (Study II).

* $p < .01$.

** $p < .001$.

† This difference is in a direction opposite to the other comparisons.

It is also possible to derive measures of the reliability of this technique from the present data. A matrix of the intercorrelations of scores for all fingers used in the first study is presented in Table 2a. Examination of this table reveals that the reliability of simultaneous measurement with different fingers uncontaminated by previous use (A-2, A-3) is acceptably high (.86). This is comparable to Brutten's (2) correlations of .92 and .93. However, with the inclusion of time intervals between measures, the reliability coefficients dropped considerably. In the first study, measures taken 10 minutes apart showed an average r of .62, those taken 20 minutes apart an average r of .72, and those 30 minutes apart an average r of .50. (Correlation coefficients were averaged by Fisher's r-to-z transformation technique.) The fact that these are both averaged and correlated correlations precludes the use of an adequate test of significance of the differences between these and the simultaneous measures, but t-tests for the difference between correlated correlations (8) run between each of the four measures separated by 30 minutes (occasions A-D) and the correlation obtained through simultaneous measurement (occasion A) show a decrease in reliability significant at beyond the .05 level for each test made. Obviously, although simultaneous measurement provides a measure of the *interchangeability* of fingers, it is not adequate as a measure of the stability of the palmar-sweat index through time.

TABLE 2A

MATRIX OF INTERCORRELATIONS AMONG PALMAR-SWEAT MEASUREMENTS TAKEN FROM THE SAME FINGER, AND FROM DIFFERENT FINGERS, ON FOUR OCCASIONS OF MEASUREMENT. LETTERS REFER TO OCCASIONS OF MEASUREMENT, AND NUMBERS TO THE FINGER USED

	A-2	A-3	B-2	B-4	C-2	C-1	D-2	D-Thumb
A-2		.86	.69	.66	.65	.70	.56	.44
A-3			.68	.67	.68	.79	.55	.56
B-2				.74	.65	.72	.58	.75
B-4					.57	.61	.32	.65
C-2						.57	.47	.55
C-1							.45	.52
D-2								.59
D-Thumb								

Reliability data from the second study are presented in Table 2b. Except for coefficients in parentheses, the data of Table 2b represent measurements made only on previously unused fingers. One would expect these correlations to be somewhat lower than those derived from the first study, since measurements from five different fingers are involved in each correlation. The fact that these correlation coefficients are not markedly lower than those reported in the first study, combined with the observation that they do not decrease significantly over periods of time up to 30 minutes, suggests a reliability among fingers used interchangeably which would be sufficiently high for most psychological and psychophysiological research in which successive measurements are taken over brief time intervals.

TABLE 2B

MATRIX OF INTERCORRELATIONS AMONG PALMAR-SWEAT MEASUREMENTS IN THE SECOND STUDY

	A	B	C	D
A	.74	.77	.76	.69
B		(.38)	.67	.81
C			(.72)	.56
D				(.75)

From the data of the second study it is possible to shed some light on the question of geographical differences among the fingers in amount of sweating produced. Since each subject produced one palmar-sweat blot from each of his five fingers which was uncontaminated by previous measurements, it was possible to compare the mean densitometer readings of blots taken from all five fingers over the 20 subjects. These means were: thumb, 7.00 ($SD = 2.31$); finger 1, 7.68 ($SD = 1.79$); finger 2, 7.22 ($SD = 2.02$); finger 3, 6.62 ($SD = 2.16$); finger 4, 6.98 ($SD = 1.83$). An analysis of variance performed on these data yielded a nonsignificant F-ratio of 0.72, offering no support for the possibility that the different fingers produce differential amounts of sweating. Since these two studies have addressed themselves only to the measurement of sweating under resting conditions, the possibility of differential reactions among the fingers to experimentally-imposed arousal situations has not been investigated.

D. DISCUSSION

The hypothesis that successive use of the same finger for palmar-sweat measurements over short time intervals yields progressively lighter blots has received unequivocal support. Significant mean differences occur between previously used and previously unused fingers on each occasion of measurement following the first in the first study and on each occasion of measurement following the second in the second study, and the mean difference between scores for previously used and previously unused fingers over time is also significant. High interfinger reliability, coupled with this lightening effect of successive measurement from the same finger makes it desirable to use different fingers for palmar-sweat measurements if the measurements are to be made in the same experimental session. The general lightening effect observed in the constant finger cannot be attributed solely to central adaptation, since the decrease in measurable sweating was significantly greater than in previously unused fingers measured *simultaneously* with the constant finger. An important factor to consider in the interpretation of these data is the apparent stability of the measurement technique itself as evidenced by the correlations of .86 and .74 in the present studies and Brutten's (2) correlations of .92 and .93 for simultaneously obtained measures uncontaminated by successive use. These measures in combination with the lower correlations obtained from measures separated in time, suggest that the variability of sweat measures is largely a function of variability of the process itself, not variability of the measuring technique.

The present studies have yielded no support for the assumption that diffeernt fingers produce different amounts of sweating at the same time.

E. Summary and Conclusions

In two separate studies, palmar-sweat measurements were taken by the colorimetric method on four different occasions separated by 10-minute intervals. On each occasion of measurement two fingers were measured simultaneously: for each subject, one finger was measured four times, each time with a different (previously unused) companion finger. It was found in both studies that successive measurement from the same finger over short time intervals introduces a measurement error which is present to a much smaller extent when each successive measurement is made from a previously unused finger. The own-control design of both studies precludes explanation of this phenomenon on the basis of physiological adaptation. It was also concluded that reliability of the colorimetric palmar-sweat technique itself is adequate for most psychological research involving repeated palmar-sweat measurements over short time intervals.

References

1. Beam, J. C. Serial learning and conditioning under real-life stress. *J. Abn. & Soc. Psychol.*, 1955, **51**, 543-551.
2. Brutten, E. J. Colorimetric measurement of anxiety: A clinical and experimental procedure. *Speech Monog.*, 1959, **26**, 282-287.
3. Davis, R. H. A further study of the effect of stress on palmar prints. *J. Abn. & Soc. Psychol.*, 1957, **55**, 132.
4. Haywood, H. C. Relationships among anxiety, seeking of novel stimuli, and level of unassimilated percepts. *J. Personal.*, 1961, **29**, 105-114.
5. ———. Novelty-seeking behavior as a function of manifest anxiety and physiological arousal. *J. Personal.*, 1962, **30**, 63-74.

6. Lawson, E. D. Attitude shift as related to palmar sweating in group discussion. Ph.D. dissertation, Univ. Illinois, 1954.
7. Light, B. H. Tension changes in patients undergoing psychotherapy. Ph.D. dissertation, Univ. Illinois, 1951.
8. McNemar, Q. Psychological Statistics. New York: Wiley, 1955.
9. Mowrer, O. H. (*Ed.*) Psychotherapy: Theory and Research. New York: Ronald, 1953.

George Peabody College
Nashville, Tennessee

Southern Illinois University
Carbondale, Illinois

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 371-385.

EXPERIMENTAL EVIDENCE FOR THE INVOLVEMENT OF A NEUROPHYSIOLOGICAL MECHANISM IN THE DISCRIMINATION OF DURATION*[1]

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND JAMES F. JORDAN

A. INTRODUCTION

Previous experiments attempting to tie biological mechanisms to duration discrimination related the latter either to rate of metabolism or autochthonous nervous activity (9, 10, 11, 12, 13, 14).

In the studies to be reported here, durations were judged when *input* to the visual system was such as to manipulate nervous activity within the optic pathway. The experiments were undertaken to discover whether nervous processes induced in the optic pathway by end organ stimulation play a role in the discrimination of duration. If such optic pathway activity is involved in the discrimination of duration of visual events, then equal intervals filled with different optic pathway activity could result in different duration discriminations.

Short photic intervals filled with intermittent pulses were used as stimuli. Each condition of intermittency (filling) was deemed capable of uniquely affecting nervous processes within the optic pathway. Stimulus conditions were decided upon by considering the neurophysiological facts and sensory phenomena underlying Bartley's "alternation of response theory."

1. *Research Orientation: The Alternation of Response Theory*

The alternation of response theory is a core framework which enables one to relate many visual phenomena to the neurophysiological activity of the visual system. It is not concerned with principles governing simply the strength and such characteristics of the afferent input from eye to the succeeding portions of the optic pathway, but is directed toward discussing the way in which such inputs are organized within the optic pathway and how this organization plays its part in determining the visual sensory end-

---

* Received in the Editorial Office on December 15, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present study is part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

result. Originally the theory was laid down to explain brightness enhancement phenomena (the Brücke-Bartley effect), but the theory may be expected to have implications for all features of visual discrimination. All inputs have a temporal aspect. The question is whether different temporal distributions within inputs are effective in helping to determine many or all aspects of sensory outcome. This, of course, would include the sensory appreciation of duration as well as the more commonly studied aspects. If this be the case, it certainly is mandatory that the appreciation of duration be studied in connection with the control of the temporal features of input.

Bartley (4) conceives the portions of the optic pathway up to and including the projection areas to consist essentially of a fixed number of parallel channels. This idea is most clearly recognizable in that section of the pathway comprised of the optic nerve where each nerve fiber may be taken to be a segment of one channel. Within the channels of the optic pathway, timing of input is all important (3). When many elements act synchronously they may summate and the extent of summation (or its opposite, inhibition) is indicated in the neurophysiological record as amplitude of wave potential. The nature of temporal distribution may therefore be described by two closely related concepts, *amplitude* and *synchrony*. Obviously if one holds the number of elements acting constant, amplitude becomes a function of synchrony.

Maximal amplitude and synchrony of individual channels are obtained when individual units in a train of intermittent inputs are too short to reactivate any of the channels, so strong that each activates all available channels, and when each input is followed by a null interval sufficiently long to permit full recovery of each channel. High levels of synchrony and amplitude are correlated with a number of paradoxical phenomena involving brightness, flicker-fusion (7), visual acuity (8), and color (6, 17). While perfect synchrony and maximal amplitude are conditions never fully realized under conditions of intermittent stimulation, they may be approached. A pulse-to-cycle fraction (PCF[2]) of .25 combined with intermittency rates around 10 cps at moderate to high levels of intensity is such a condition (2, 3, 5).

When duration of the photic pulse in a given intermittency cycle is lengthened, amplitude is reduced through the reduction of synchrony. This occurs because photic pulses of longer duration reactivate single channels and leave too little time for recovery of the last channels activated before the onset of the next pulse. As the pulses are lengthened they come to act more like continuous illumination. Minimal synchrony is produced under conditions of prolonged uniform stimulation. Synchrony is obliterated because recovery time varies somewhat from channel to channel and even after a short exposure it may be expected that there will be as many channels going into action as going into recovery phase at any instant. Uniform activity throughout time is to be expected, the amplitude of the activity depending upon the intensity of the impingement. In general then, beyond a critical intensity with all else equal, the longer the pulses in an intermittent photic train, the greater the desynchronization and the lower the amplitude until full alternation of channel response is achieved.

---

[2] In older terminology this is light-dark-ratio.

Effects from other conditions are also predictable. A PCF .75 under circumstances otherwise identical to those outlined will produce a lesser effect than PCF .25 in all respects, but a greater effect than steady stimulation. However if rate of photic intermittency is sufficiently increased, effectiveness occurs in the opposite order. The photic period ("light" period) of the larger PCF becomes more nearly optimal than that of the smaller PCF even though the null period ("dark" period) separating the photic intervals becomes increasingly less adequate for full recovery. The null photic interval is obviously of lesser importance than the longer photic intervals when rates are high. A PCF of .75 is three times as bright as a PCF of .25 at high rates of intermittency, e.g., 50 cps; and a steady target, a condition devoid of a null period, is four times as bright as a PCF of .25 (Talbot-Plateau Law).

B. First Study

An exploratory study aimed to discover whether the mechanism conceptualized in the alternation of response theory could be demonstrated to be utilized in the discrimination of duration. It was reasoned that an effect would be demonstrated if intervals of equal clock length but differing in photic intermittency could be shown to be differently discriminated as durations.

Specifically it was expected that: (*a*) a PCF .125 at 10 cps at an intensity level producing brightness enhancement would result in either higher or lower duration discrimination than PCF .125 at 54, 15, or 2.5 cps and PCF .875 at 54, 15, 10 or 2.5 cps since it is associated with higher order amplitude and synchrony of discharge in the optic pathway than are any of the latter seven intermittency conditions; (*b*) discriminations associated with PCF .125 at 2.5 and 15 cps would tend to be of a magnitude intermediate between those under PCF .125 at 10 and at 54 cps since these produce brightness levels lying between maximum enhancement level (PCF .125 at 10 cps) and the Talbot level (PCF .125 at 54 cps) which are the extremes of brightness associated with intermittently presented illumination at a given intensity level; (*c*) duration discriminations made under PCF .875 condition at any of the rates of intermittency used should deviate from PCF .125 at 10 cps in the same direction as those of PCF .125 at 2.5, 15, and 54 cps, because PCF .875 will not produce brightness enhancement at any rate of intermittency under ordinary conditions of intensity; and (*d*) the index of duration discrimination should vary slightly but progressively from PCF .875 at 2.5 cps to 54 cps because increasing rate reduces brightness, this progressive increase or decrease in the direction of the difference between PCF .125 at 10 cps and PCF .875 in general.

1. *Method*

A Grason-Stadler electronic circuit driving a glow modulator tube provided control of PCF, rate, and intensity. Since the luminous element of the glow modulator tube is small, a biconvex lens was installed to increase the effective area. The lens producing this Maxwellian view was stopped to give a target of 2° visual angle from a chin rest 15 inches away.

Five males and five females, all students in an introductory psychology class, served as subjects. Each served singly. The experiment was conducted in a dark room. The subject was seated and the chin rest adjusted for use of the right eye. The left eye was occluded. The instructions were to fixate upon the lens, note the interval during which it was illuminated, and to reproduce its duration with an auditory continuum immediately afterward by pressing a button actuating a buzzer. The subject was instructed to form his discriminations on a direct and not an indirect basis. That is, he was specifically asked not to use self-produced measuring devices such as tapping, counting, etc. Several practice trials were given using a continuous photic target and employing intervals differing in magnitude from those used during subsequent actual experimentation.

During the experiment proper, presentations of .5, 9, 3.0 and 7.0 secs. were discriminated 10 times each under each of the eight conditions of intermittency. The total 320 trials were completely randomized for each subject, and only 80 trials were completed in a session. Intervals were produced by placing an electrically timed shutter in front of the luminous element of the glow tube. The tube was made to discharge in the desired fashion before the shutter was activated. All of the above devices operated silently.

Intensity was set at 3,240 c/ft^2 for all conditions. Thus, while all targets were of the same intensity, total flux per cycle varied. That is, PCF .875 delivered per cycle seven times as much energy to the eye as PCF .125 even though when "on" the presentations were equally intense.

Several features of the experimentation must be noted. *First*, since the target was a Maxwellian view, it did not work so well for the briefest intervals presented. No way was found to hold the fixation and slight eye movements were critical in impoverishing the view. Nevertheless the set-up would have been satisfactory for the purposes at hand had not there been long null photic periods between presentations. Therefore when later inspection of the data showed a large response variation associated with various blocks of short presentations it was decided to consider only the upper end of these response distributions, assuming that for such responses the presentations were seen in their entirety. While this selection does not bias results, it does not permit comparing short presentations as regards overestimation or underestimation.

Second, intensity proved so great as to provoke discomfort. Some subjects reported difficulty in fixating the target under enhancement conditions.

Third, the rate of 2.5 cps probably involved the accommodative mechanism. Some accommodation occurs below five cps (2) and discomfort was occasionally reported by subjects at this rate.

Fourth, and perhaps most disturbing, was the use of self-produced standards. It proved difficult to convey to the subject that the "duration" involved in this study was a personal appreciation or response rather than something external to the subject, such as clock time. Despite instruction, virtually every subject reported using devices such as counting, tapping, etc. during the experiment. In order to assess the effects of employing self-produced standards, one would have to know the effects of the intermittency variables on tapping rates. Certainly Hoagland's experiment (11) showed a relation between autochthonous cortical rhythms and tapping rate, and LeGrand (16), Bartley (1), and Segal (18) each demonstrated that the discrimination of flicker rate is linked to brightness discrimination. Perhaps the method of reproduction affords too great a temptation and ought to be abandoned except for very short intervals.

It seems very likely however that the effect of self-produced standards was less for the short than for the long intervals. The senior author's observations indicate that a .5-second interval is too short to provide opportunity to use self-produced standards, a .9-second interval provides little opportunity, 3.00 seconds a distinct opportunity and 7.00 seconds a good opportunity. The progressive change in curves from one duration to another seems to bear this out, as shall be seen.

2. *Results*

The data obtained are shown in Figure 1. As said, the curves representing the .5 and .9 intervals (Figures 1, 1 and 2) were determined by taking the mean of the two largest responses given by each subject under each condition. Each of these points therefore represents the mean of 20 observations. For these two conditions the mean magnitude of matchings for PCF .125 at 10 cps is greater than for any other condition. This is in accord with expectations. Matchings of PCF .125 at 15 cps fall between points representing 10 and 54 cps as it was supposed they would. PCF .125 at 2.5 cps does not meet expectations quite fully. It is ranked below 10 cps rather than between 10 and 54 cps. There were no noteworthy exceptions to expectations regarding PCF .875. All responses are in the expected relation to PCF .125 at 10 cps and a decline with rate is observable.

In all cases, the differences alluded to are small, but such as they are, they are largely in the expected direction. Statistical test of the differences is difficult. However, the two graphs show the high points of the curves fall at PCF .125 at 10 cps and the low at PCF .875 at 54 cps. The individual data of nine of 10 subjects making up Curve 1 show this while the data of the tenth subject show virtually no difference between conditions. For Curve 2, the data from all 10 subjects are entirely consistent with the points graphed.

Curves 3 and 4 of Figure 1 represent the full data for the 3.0- and 7.0-seconds interval conditions with each point representing 100 observations. These curves tell another story. In fact, one can see a progressive shift in function going from Curve 1 to 2 to 3 to 4. The curve for PCF .125 in 4 is virtually the inverse of that of 1 and the same thing holds true for the PCF .125 Curves for 2 and 3. Also, PCF .875 shows a constant increase in steepness of slope from 1 to 4. Moreover there is a gradual shift in position. In Curve 1 all points of PCF .875 are below those of PCF .125. This is true for all but one point in Curve 2, two points in Curve 3, and in Curve 4, all but one point. However the results for individual subjects are much less homogeneous than for the shorter intervals. The data of six of the 10 subjects making up Curve 3 fit the plots of 1 and 2 more closely than those of 3 and 4. This is also true of the data determining Curve 4. This and the progressive shift in effect suggests involvement of a factor additional to stimulus intermittency. An increasing involvement of self-produced standards in the judgments as interval length is increased is one possible explanation.

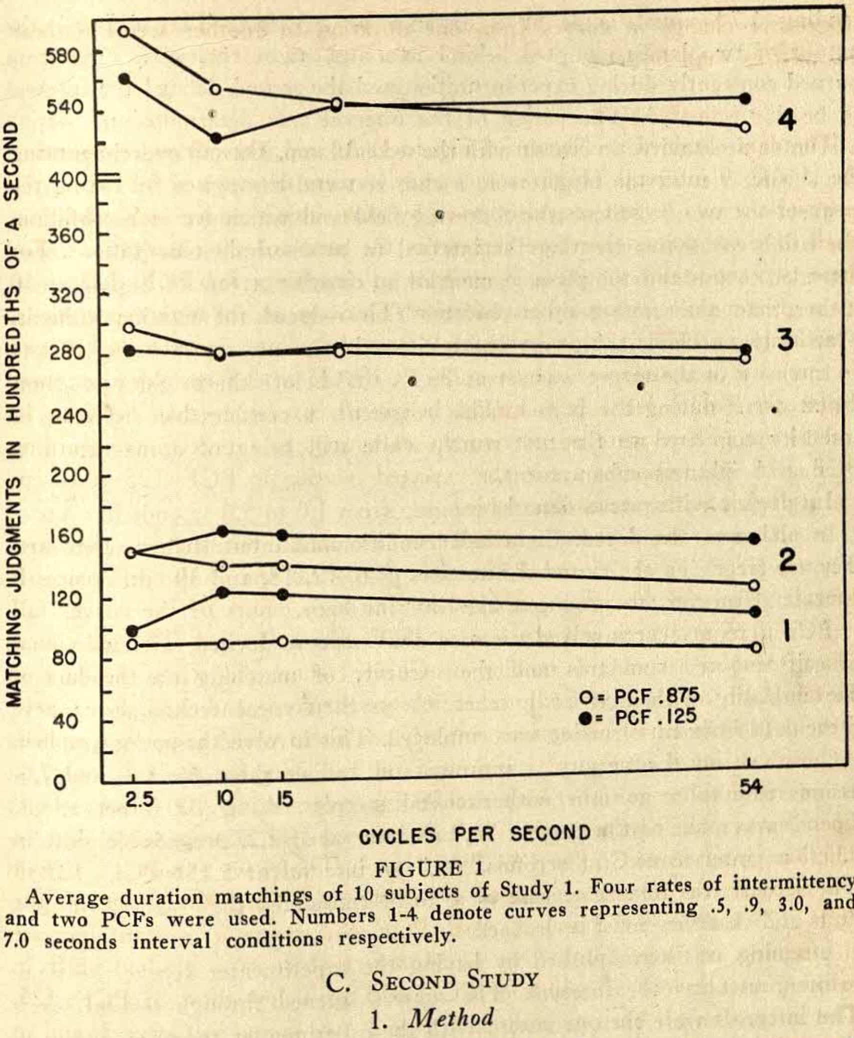


FIGURE 1
Average duration matchings of 10 subjects of Study 1. Four rates of intermittency and two PCFs were used. Numbers 1-4 denote curves representing .5, .9, 3.0, and 7.0 seconds interval conditions respectively.

C. Second Study

1. *Method*

In a similar study 12 students in an advanced undergraduate experimental psychology class served as subjects. Experimentation was preceded by lectures dealing with optic pathway mechanism and by a number of hours observing brightness and flicker-fusion phenomena. The level of understanding was low, and not connected with the experimental discrimination required.

The Grason-Stadler was abandoned and the problem involving Maxwellian view thereby eliminated. The target was a square opal glass surface subtending 1.7° visual angle at a distance of 5.5 feet. This diffused the output of two lamps mounted behind in a light-tight enclosure. One lamp burned constantly during experimentation and the second defined the interval to be discriminated. The period of the interval was determined by setting a Hunter timer wired in circuit with the second lamp. During experimentation the subject viewed the target from a chin rest and through a field stop. His instructions were to fixate the flickering field and watch for it to brighten, the bright phase representing the interval to be discriminated. The interval attended to was thus simply a segment of an ongoing stream of impingements rather than a discrete train as before. This reduced the fixation difficulty previously associated with very short intervals.

Intensity of the target was set at 76.7 c/ft.^{2} before the bright period and 140.2 c/ft.^{2} during the period. This represents a considerable reduction in intensity compared to the first study, while still being of a magnitude to produce brightness enhancement.

Intervals of 21 separate lengths ranging from 1.0 to 5.0 seconds in units of .2 seconds were employed. Eight new conditions of intermittency were used. These were PCFs of .25 and .75 at rates of 6, 12, 18, and 30 cps. A silently operating episcotister produced the intermittency.

A final refinement involved use of a new response devised to eliminate use of self-produced standards and the necessity of matching the product of one modality with that of another. A psychophysical technique used by Stevens (19) in ratio scaling was employed. This involved assigning numbers to positions on the sensory continuum and having the subject use this to assign magnitudes to other experimental samples. A steady target of .83 seconds was assigned the number "1," an interval of 2.27 seconds the number "35," an interval of 3.71 seconds "69," an interval of 5.15 seconds "103." The subject was told that these were to be considered the "names" of these units and that they must be learned.

Learning was accomplished by having the experimenter repeat the name before presenting the interval. The subject listened through several series. The intervals were then presented with the experimenter repeating the name after the interval and the subject repeating the name after the experimenter. Finally the subject was asked to name them unassisted. The learning criterion was two successive correct namings of a random series. The criterion met, the subject was told that if he would consider the number series to parallel his appreciation of duration, he now had a name for any duration discrimination he might make within these limits.

A few subjects initially asked if they might count or tap. This was forbidden. It was explained that many intervals would be too short for counting or tapping, and that reckoning from some such standard in any exact way would require the use of decimals since all the names learned were odd numbers. In practice, all found the four standard intervals clearly distinct from one another and none of the subjects found use of self-produced standards necessary. It may be said that every subject understood the procedure and acknowledged capability to use the device.

Each subject served under all eight conditions of intermittency. These were randomized differently for each subject. The 21 intervals within each condition were presented in random order. Before presenting an interval the experimenter called out "ready." The subject was instructed to respond immediately after each presentation.

2. *Results*

Figure 2A plots the percentage deviations of the means of each intermittency condition from the average interval length (3.0 seconds). Individual data

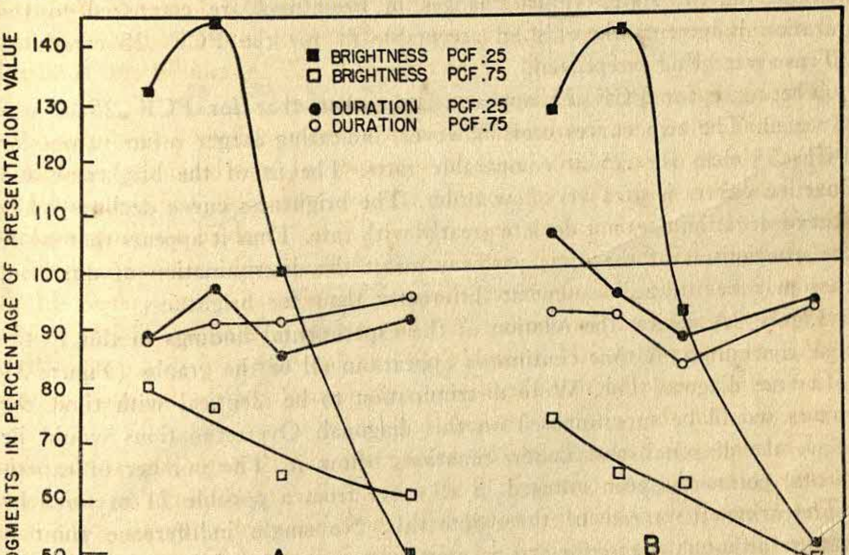


FIGURE 2

Average duration matchings for subjects of Study 2 and Study 3 are given in A and B parts respectively. Brightness matchings made under corresponding conditions by the senior author are included for purposes of comparison. Effects from five rates of intermittency and two PCFs are compared in each case. All matchings are given as percentage of standard condition.

were consistent with the means plotted. The high and low points for 11 of 12 subjects occurred in the same relationship as in Figure 2. Statistical test of these differences is difficult for the same resasons as in Experiment 1.

The graph shows PCF .25 at 12 cps to again produce mean judgments of largest magnitude. Insofar as rates of intermittency overlap, the PCF .25 curve is similar to 1 and 2 plots of Figure 1 for PCF .125, the chief exception being the 30 cps condition. Also when the means of intervals 1.0 second through 1.4 seconds, 1.6 seconds through 2.0 seconds, . . ., and 4.6 seconds through 5.0 seconds, were plotted it became evident that the curve form holds true for lesser segments of the interval range as well.

Figure 2A also allows comparison between changes in target brightness and duration. Brightness is plotted in terms of per cent deviation from the brightness value of the steady target. The upper (PCF .25) brightness curve shows the familiar Brücke-Bartley effect or brightness enhancement. At low rates of intermittency the target appeared brighter than the same target in a non-intermittent state. When changes in brightness are compared to the duration judgments there is an acceptable fit for the PCF .25 condition, 30 cps again being exceptional.

The curve for PCF .75 appears flatter than that for PCF .25 as was expected. The two curves cross however, indicating larger mean values for PCF .75 than for .25 at comparable rates. The fit of the brightness and duration curves is of a very low order. The brightness curve declines while that of duration does not deviate greatly with rate. Thus it appears that while the mechanisms of the optic pathway affect the discrimination of duration, they may be utilized somewhat differently than for brightness.

Figure 3A depicts the relation of the experimental findings in this to the time continuum. A time continuum appears in all of the graphs (Figure 3) as a true diagonal line. Were discrimination to be identical with time, the curves would be superimposed on this diagonal. Overestimations would lie above the diagonal, and underestimations below it. The number of experimental points has been reduced in all cases from a possible 21 to seven by representing the mean of three intervals. No single indifference point is present although overestimation is most common with briefer intervals. The largest number of overestimations occurs with the brightness enhancement condition PCF .25 at 10 cps.

D. THIRD STUDY

1. *Method*

The third study utilized the same equipment and was conducted in the same manner as the second except for the following changes.

Each of 64 subjects served under two conditions, one of PCF .25 and one of PCF .75. The experiment was counterbalanced with respect to PCF and rate. Only totally naive and untrained subjects were used. Intensity was reduced

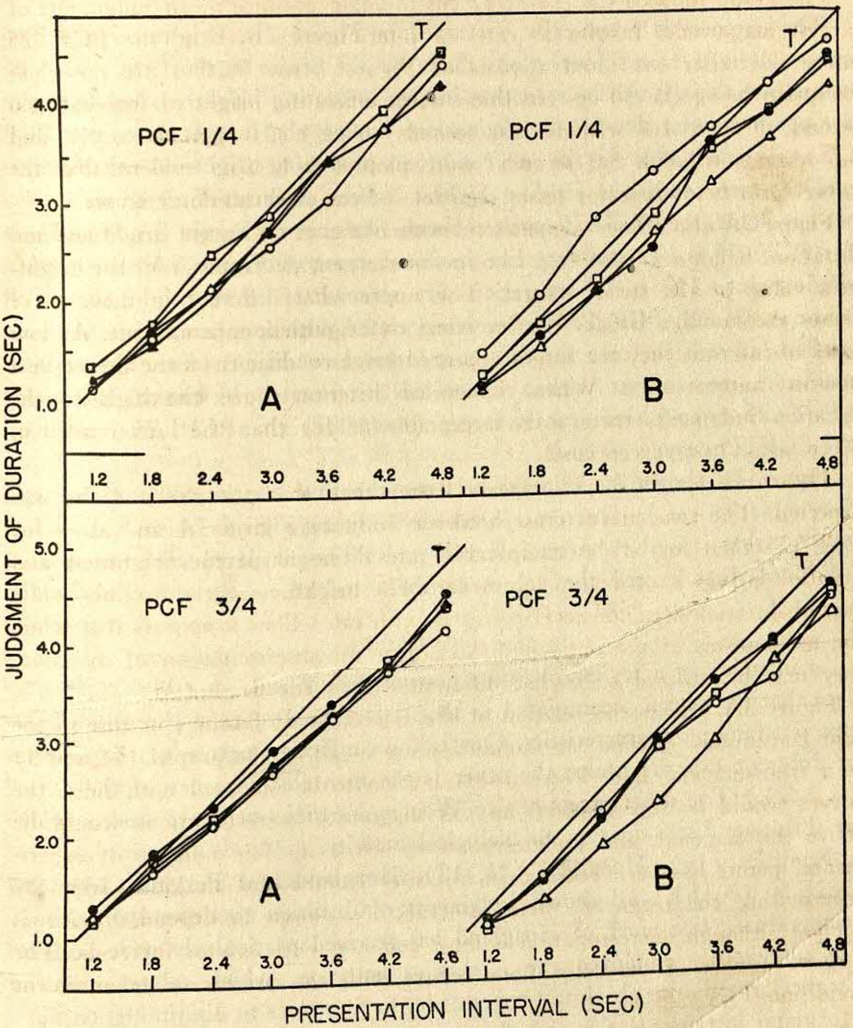


FIGURE 3

Plots showing average duration matching associated with each standard interval. Every combination of intermittency and PCF is represented. Perfect invariance between clock interval and duration judgment would lead all points to fall on line "T." Points falling above this line represent overestimation and those below, underestimation. Portions of the figure labeled A refer to data of Study 2, those of B to Study 3.

from the level in Study Two. The bright period was now 85.6 c/ft.^{2} and the illumination between intervals, 27.4 c/ft.^{2}.

2. *Results*

The outcome is graphically portrayed in Figure 2B. Brightness matching curves derived from observations made by the senior author are given. By comparing these it can be seen that indices indicating brightness enhancement are again associated with discriminations having the longest duration. PCF .25 at six and 12 cps are the two conditions producing brightness enhancement and the two conditions yielding highest indices of duration. Except for the six cps point, the PCF .25 curve is much like that of Figure 2A. The curve for PCF .75 is likewise very like its counterpart of Figure 2A except for a single point. The curve separations are somewhat different in 2A and 2B however. In 2B, PCF .25 is always associated with longer duration discriminations, as was the case for the short interval conditions of the first experiments. A median test with a one-tailed criterion shows the highest point (PCF .25, 6 cps) significantly larger ($p < .001$) than the lowest point on the PCF .75 curve (30 cps).

Figure 3B shows the relation between the time continuum and duration discrimination. It has been prepared similarly to Figure 3A and shows, in general, that it is the shorter intervals and those associated with paradoxical brightness that exceed the values of the standard.

E. Discussion

In introducing a review of "time perception," Woodrow (20) said: "The data that have been accumulated in the illusive field of time perception show two outstanding characteristics. One is the conflicting nature of the findings of different experimenters; the other is the mentalistic nature of the data." (p. 1224.) The main exceptions to Woodrow's indictment are studies which tie duration discrimination to biological activity.

Francous (10), Hoagland (11, 12), Hoagland and Perkins (13), and DuNoüy (9) have all shown judgments of duration to depend upon body temperature. The work of Hoagland has received particularly wide acclaim. He successfully employed the Arrhenius equation, which relates speed of chemical reaction to absolute temperature, to fit curves to discrimination data. Hoagland interpreted this as reflecting the control of autochthonous activity within the brain by a chemical pacemaker.

Stimulus features were not manipulated in these older studies. Thus only a restricted sample of the possible pool of effective biological mechanisms was touched upon. Of those omitted, the role of afferent activity is of particular interest since most discrimination of duration does involve external events. A step toward elucidation of this problem has been made by Holubar (14)[3] who studied the relation of intermittent photic input to duration. He suggests such discrimination has a phase relation to alpha rhythm.

The present study also involved intermittent photic stimulations and durations. This was the means used to connect visual discrimination of duration to amplitude and synchrony of activity in the optic pathway.

The results indicate such a relationship is present, for two reasons. First, photic cycling likely to increase synchrony and amplitude produced the longest duration discriminations. Second, the largest mean indexes of duration were found always associated with conditions in which amplitude and synchrony were expected to be near maxima. The opposite was not always true however, since high synchrony and amplitude did not always produce the largest indexes of duration. It is possible that correspondence might have been improved by keeping the phase relation between rate of repetitive input and alpha rhythm optimal. The short interval curves for Study 1 (Figure 1, Curves 1 and 2) suggest this since they both represent more perfect phasing and lack some of the irregularities present in the data from the later two experiments (Figure 2, A and B). Also, the 30 cps points of Graphs 2A and B both represent good phasing and are unexpectedly high. Since the human alpha rhythm is near 10 per second, 30 of every 90 pulses would be in phase at 30 cps, presuming cortical driving to be possible, i.e., displacement of the alpha rhythm by means of repetitive afferent input (5). The lowest point on the other hand represents the poorest phasing. A rate of 18 cps is in phase with alpha rhythm only five times every 90 pulses. Jasper (15) has in fact reported "desynchronization" of cortical waves to occur in the neighborhood of 18 cps. The chief difficulty with this interpretation is that the brightness curves were not so affected as one would expect if the above were true. There is however, a possibility that the forementioned interaction is of consequence only during the relatively short periods represented by the intervals discriminated for duration.

The duration continua for intervals filled with photic intermittencies were all essentially linear and parallel to one another although not precisely parallel to the time continuum. It may be inferred that the specific location of the "indifference intervals" have to do with both the range of intervals used, i.e., show a series or adaptation level effect, and, importantly, the specific nature of the photic presentation. While indifference points always occur, the more effective the condition in extending duration, the longer the indifference interval. When one recalls that the second and third designs separated the experimental conditions from one another, and that in the third study each individual saw only two conditions in total, it is not possible to account for specific indifference points by the concept of adaptation level alone.

---

[3] Seen in abstract only.

F. SUMMARY

In three experiments, naive subjects discriminated the duration of intervals filled with intermittent photic stimuli. The photic presentations selected were such as to vary amplitude and synchrony of optic pathway activity. The results indicate that (*a*) amplitude and synchrony of optic pathway activity do have an effect although this effect perhaps interacts with autochthonous cortical activity; and (*b*) the continuum of duration discrimination is linear but not precisely parallel to the time continuum, the indifference point depending upon adaptation level and the specific nature of interval filling. There is some evidence indicating that self-produced standards are used when the experimental design makes this possible and are likely an important variable in discrimination of longer intervals. A methodology was devised to eliminate this practice.

REFERENCES

1. BARTLEY, S. H. Subjective flicker rate with reference to critical flicker frequency. *J. Exp. Psychol.*, 1938, **22**, 388-394.
2. ———. The psychophysiology of vision. *In* Stevens, S. S. (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951.
3. ———. Central mechanisms in vision. *In* Field, J., Magoun, H. W., & Hall, V. G. (*Eds.*), *Handbook of Physiology, Sec. 1, Neurophysiology, Vol. I.* Washington, D.C.: Amer. Physiol. Soc., 1959. (Chapter 30, pp. 713-741.)
4. ———. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Ophthalmol.*, 1958, **60**, 775-791.
5. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. *In* Jung, R., & Kornhuber, H. (*Eds.*), *The Visual System: Neurophysiology and Psychophysics*, Heidelberg: Springer-Verlag, 1961.
6. BARTLEY, S. H., & NELSON, T. M. Certain chromatic and brightness changes associated with rate of intermittency of photic stimulation. *J. of Psychol.*, 1960, **50**, 323-332.
7. ———. A further study of pulse-to-cycle fraction and critical flicker frequency. A decisive theoretical test. *J. Opt. Soc. Amer.*, 1961, **51**, 41-45.
8. BARTLEY, S. H., NELSON, T. M., & SOULES, E. M. Visual acuity under conditions of intermittent illumination productive of paradoxical brightness. *J. of Psychol.*, 1963, **55**, 121-143.
9. DUNOÜY, R. Biological Time. New York: MacMillan, 1937.

10. FRANCOUS, M. Contribution a l'étude du sens temps. La témperature interne comme facteur de variation de l'ápreciation subjective des durées. *Année Psychol.*, 1928, **28**, 186-204.
11. HOAGLAND, H. The physiological control of judgments of duration: Evidence for a chemical clock. *J. Gen. Psychol.*, 1933, **9**, 267-287.
12. ———. Pacemakers in Relation to Aspects of Behavior. New York: Macmillan, 1935.
13. HOAGLAND, H., & PERKINS, C. T. Some temperature characteristics in man. *J. Gen. Physiol.*, 1935, **18**, 399-408.
14. HOLUBAR, J. The time sense and photic driving in man. *Physiol. Bohemaslov.*, 1960, **9**, 482-487.
15. JASPER, H. H. Reports and communications. *XI Internat. Congr. Psychol.*, 1937, 226.
16. LE GRAND, Y. Sur le rhythme apparent du papillatement. *C.R. Acad. Sci.*, 1937, **204**, 1590.
17. NELSON, T. M., & BARTLEY, S. H. The role of PCF in temporal manipulations of color. *J. of Psychol.*, 1961, **52**, 457-477.
18. SÉGAL, J. Le méchanisme de la vision en lumière intermittente. Paris: Alcon, 1940.
19. STEVENS, S. S. To honor Fechner and repeal his law. *Science*, 1961, **133**, 80-86.
20. WOODROW, H. Time perception. *In* Stevens, S. S. (*Ed.*), *Handbook of Experimental Psychology*. New York: Wiley, 1951.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 387-396.

THE BRIGHTNESS OF PART-SPECTRUM TARGETS STIMULATING THE EYE PARAFOVEALLY AT INTERMITTENCY RATES BELOW FUSION*[1]

Department of Psychology, Michigan State University

THOMAS M. NELSON, S. HOWARD BARTLEY, AND ZACHARY FORD

A. PROBLEM

Bartley and Nelson (4) and Nelson and Bartley (18) recently described color phenomena they designated by the terms "washout effect" or "color bleaching." These terms refer to a phenomena-complex wherein a simultaneous loss of saturation and shift in hue occurs when a part-spectrum target-field is presented to the eye under special conditions of stimulus intermittency. Brightness changes accompany the hue-saturation changes. Within a certain range productive of brightness enhancement an intermittent target-field is seen as brighter than the steady target-field whereas one would likely expect it to lie somewhere between the steady and Talbot-level in brightness. They found the magnitude of this paradoxical brightness larger with some wavelength distributions than others and larger with low rates of intermittency and small PCFs (pulse-to-cycle fraction[2]), than the reverse.

Quantitative data on the part-spectrum effects have not been published even though the presence of the phenomenon has been established using objective criteria (21). The present paper reports quantitative data on the brightness factor in washout effect. In the study to be reported five part-spectrum distributions were made intermittent and matched for brightness with a steady target passing the same wavelength.

B. METHOD

1. *Apparatus*

Equipment consisted of a chin rest, fieldstop, episcotister driven by a synchronous motor, a double lamp housing, two auto-transformers, a log .5 neutral density filter, and pairs of numbers 70, 72b, 73, 74, and 75

---

* Received in the Editorial Office on December 16, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The present paper is a part of the work being done under Research Grant NSF-G-19485 from the National Science Foundation.

[2] In older terminology this is light-dark-ratio.

Wratten Color Filters. Filter 70 has peak transmission near 700 mμ, 72b near 600 mμ, 73 near 570 mμ, 74 near 530 mμ, and 75 near 490 mμ. The range of transmission is 50 mμ, 60 mμ, 50mμ, 70 mμ, and 80 mμ, respectively. Complete specifications are given in the Eastman catalogue.

Both sides of the double lamphousing contained G.E. "artificial daylight" filament lamps. A rectangular opening was placed in front of each lamp. The right served as the experimental, the left as the matching target. When viewed from the chin rest and through the fieldstop both targets subtended a visual angle of 2.36° horizontally and .67° vertically. Their separation was 1.77°. The lamps for the targets were activated by 60 cycle alternating current and separately controlled in output.

The Wratten filters were placed in front of the diffusing glass but back of the rectangular openings defining the targets. The intensity of the flux through the filters was always pre-set so as to provide for targets that would be equal in luminosity when non-intermittent. Flicker photometry was used to establish these equal luminosities. All were equated to the "red" (Wratten 70) which was set to radiate 48.39 c/ft.^{2}.

An episcotister driven by a synchronous motor provided target intermittency. Discs were cut to provide PCFs of .25, .50, and .75 at rates of 5, 10, and 20 cps. Owing to their large size the discs assured photic pulses with adequately sharp onsets and terminations.

2. *Observers*

The three authors served as *O*s. Observers N and B are experienced observers. Observer F had little experience prior to the study. Totally naive *O*s were avoided because the judgment involved was difficult and variability could be expected to be great (19).

3. *Procedure*

Monocular right-eye observations were made from the chin rest as *O* looked through the fieldstop and fixated on a dot midway between the targets. Stimulation was therefore parafoveal. Each set of observations was preceded by a one-minute period of stimulation at pre-set intensities because temporally induced color phenomena are known to change as duration is extended (1, 13).

Comparisons of brightness were made using the method of limits. On the even numbered trials *O* turned intensity upward until the brightness could no longer be said to be dimmer and on the odd numbered downward until it could no longer be said to be brighter. Five ascending and five descending trials were taken under each intermittency condition to establish a mean. One color was viewed at a single sitting. Observations under all intermittency conditions accompanied each filter condition. Color and intermittency conditions were presented at random.

Unequal numbers of observations were made by each *O*. N made a large number of matchings over a period of several months. His data is the mean of all observations. The data of B are based upon one set of observations and that of F on two sets.

4. *Methodological Note*

Extensive chromatic alterations were produced under some of the intensity conditions used. It is perhaps appropriate to reiterate that two targets passing appropriate and identical portions of the spectrum, of identical intensity, will be seen as two unique colors providing one is presented as a series of pulses with rate and photic and null periods such as to produce near optimal washout effect. When conditions are appropriate, filter 70 will look to be a desaturated yellow while the matching target remains a deep red, or filter 75 will appear deep blue as a steady target and fully desaturated (white) when intermittent (4, 18). The fact that different chromas are being matched increases variability because it means, from *O*'s point of view, that despite the presence of identical wavelengths, "heterochromatic photometry" is being attempted using the "equality of brightness" technique. The impracticality of this technique is well known. Ives (15, 16, 17) long ago showed that data of different *O*s and the same *O* on different days are highly unreliable, that the magnitude of the matchings is affected in an unsystematic way by the wavelength characteristic of the referent standard, and that the greater the difference in hue the more difficult and unsatisfactory the match. We can expect all these variables to enter for certain of the conditions, particularly PCF .25 at five and 10 cps. An alternative to the present method is to make compensatory saturation and hue adjustment from condition to condition supposing that with these adjustments impingements producing similar colors will result. However the act of making these adjustments would alter the very conditions set out to be investigated and would in itself presume precise knowledge concerning washout effect not now possessed.

C. RESULTS

Figures 1 and 2 represent the data of N graphically. Brightness, PCF, rate, and wavelength are variables in both figures. Figure 1 plots brightness against rate for each filter and PCF. Figure 2 shows how brightness varied with wavelength for each PCF and rate. The brightness index represents a ratio of the mean of intermittent condition matchings to the mean intensity of the steady condition matchings, the matchings themselves being expressed in c/ft.^{2}. An index of 100 represents a match just equalling the steady; 50, a match half as intense as the steady; 200, twice as intense, etc. All points above 100 therefore indicate chromatic brightness enhancement.

With respect to brightness enhancement it is perhaps appropriate to point out that this is a sensory effect for which no suitable units have yet been developed. Instead, brightness enhancement has been dealt with only in terms of the relative intensities of the steady and intermittent targets necessary to provide a brightness match. This is an arbitrary procedure and an index of 200 should not be taken to mean that the matched intermittent target was seen to be twice as bright as a steady. The indexes merely are ordinal measurement and are not additive in the strict sense.

Returning to Figures 1 and 2 it is readily apparent that a wide range of brightness differences was produced. Matching indexes range roughly between 30 and 300. Also brightness enhancement occurs frequently. Twenty-two of the 45 points plotted exceed the index 100. Figure 1 shows rate to be a critical variable. Intermittency of five cps produced enhancement in 13 of 15 cases; 10 cps in eight of 15 cases; and 20 cps in only one of 15 cases. Index magnitudes clearly fall in the same order, the highest values being associated with five cps and the smallest with 20 cps.

The role of PCF in determining the brightness indexes is more complex since it also depends upon rate, i.e., the Talbot-Plateau law. At optimal frequencies, five and 10 cps, the order of effectiveness is PCF .25, .50, and .75 in all cases. Notice that this relation is the inverse of that expressed by the Talbot-Plateau law for fusion frequencies since the smallest of the three PCFs, and hence the smallest average quantity of photic radiation produced the greatest brightness whereas the same produced the least brightness at 40 cps, which is approximately the rate of flicker-fusion. Rate and PCF therefore operate as a single factor so far as intermittent part-spectrum brightness enhancement is concerned.

The 20 cps brightnesses apparently represent a range of frequency intermediate between enhancement and Talbot-Plateau levels. Brightness enhancement does not occur in any of the 15 cases and values are rapidly declining or already at their final value. The rate of decline appears to be somewhat dependent upon the magnitude of enhancement effect but the pattern is not quite definite enough to merit comment.

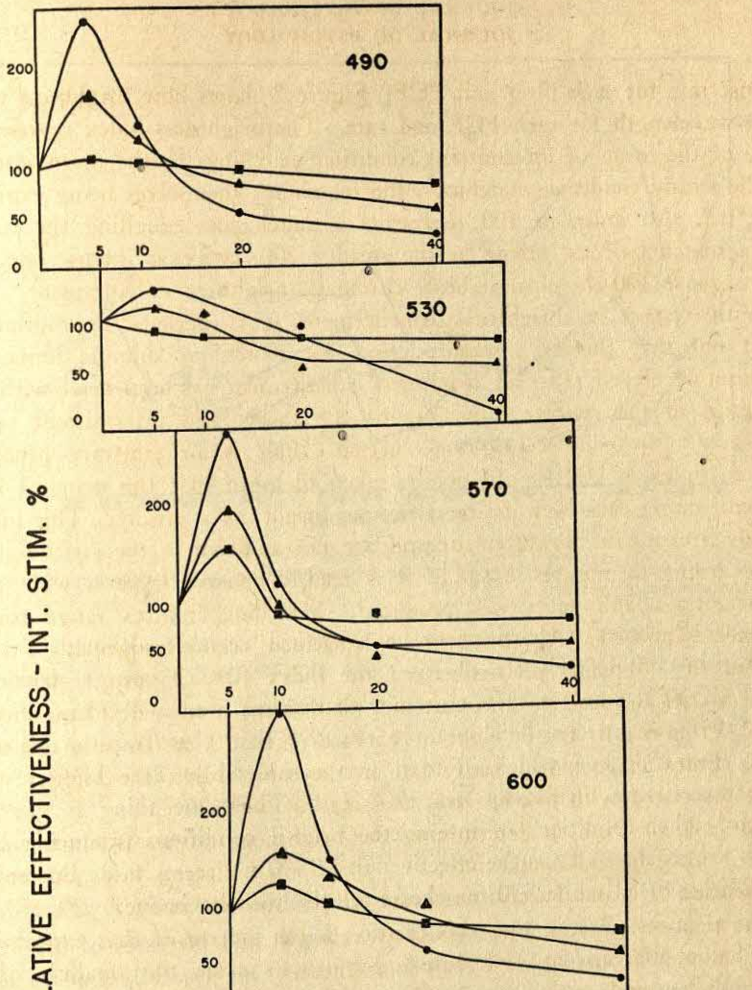


FIGURE 1
Brightness curves for observer N presented to show the role of rate and PCF in chromatic brightness. The extreme right points in each case represent Talbot-Plateau levels.

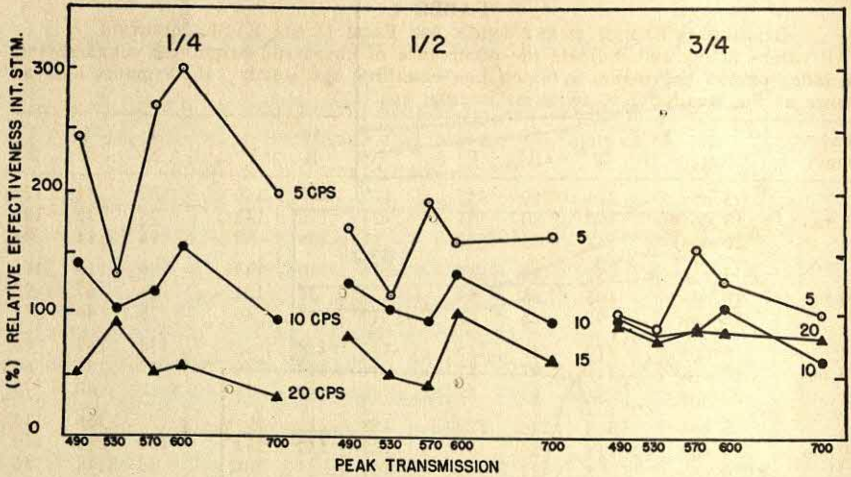


FIGURE 2

Brightness curves for observer N presented to show the role of rate and wavelength in chromatic brightness.

Figure 2 shows wavelength to be a critical variable although not one that operates independently of rate and PCF. Under optimal conditions (PCF .25 at five cps) targets transmitting maximally near 490 mμ, 600 mμ, and 570 mμ were considerably more effective than that transmitting maximally near 530 mμ which in turn was considerably more effective than that transmitting maximally near 700 mμ. The same thing is generally true for other conditions of intermittency until conditions producing lesser effects are reached. Thus the effectiveness of 490 mμ seems to decline rapidly as PCF is shifted while 530 mμ shows relatively little change.

The relation between increases of wavelength and increases of brightness, other variables constant, is clearly non-linear. One typically finds a single maximum and a single minimum value and a point of inflection somewhere between. The relation seems describable by a cubic parabola. Therefore as sheer abstract form, it is similar to the curve of Purdy (23) depicting the effects of a 10-fold increase of intensity upon hue. The peak effects from intermittently induced brightness are shifted approximately +70 mμ as compared to those occurring for Bezold-Brücke effect however.

Table 1 gives the data for all three *O*s. Absolute enhancement values are different from *O* to *O*. In addition there is less internal consistency in the tabled data of B and F than of N, presumably because of differences in the number of observations determining the mean (20), and in

TABLE 1

BRIGHTNESS ENHANCEMENT INDEX FOR EACH *O* FOR EACH CONDITION

All values above 100 indicate the occurrence of chromatic brightness enhancement. An index of 100 represents a match just equalling the steady; 50, a match half as intense as the steady; 200, twice as intense, etc.

Wave length		PCF		.25			.50			.75	
	Rate	*O*	N	B	F	N	B	F	N	B	F
490 mμ	5 cps		244	259	355	170	227	350	102	84	152
	10 cps		140	80	335	124	150	148	98	125	128
	20 cps		51	59	152	81	108	90	96	111	81
530 mμ	5 cps		132	44	167	114	49	143	89	14	116
	10 cps		103	68	64	103	30	121	84	67	92
	20 cps		92	56	47	49	44	60	79	44	71
570 mμ	5 cps		270	157	115	191	258	82	153	149	86
	10 cps		118	97	87	94	113	69	87	131	60
	20 cps		51	81	48	41	88	74	87	80	79
600 mμ	5 cps		301	176	222	158	215	196	127	156	113
	10 cps		155	124	192	132	133	191	106	215	130
	20 cps		58	33	103	101	85	106	85	118	66
700 mμ	5 cps		198	189	148	164	229	128	99	176	126
	10 cps		95	22	85	94	218	82	62	60	86
	20 cps		32	20	14	62	80	63	79	169	62

the case of F, skill in matching (19). The relative magnitude of the brightness matchings follows the same pattern, however. The most effective combinations of conditions were those involving filters 75, 73, and 72B at a rate of five cps and a PCF of .25. When these conditions were met the matching result exceeded an index of 150 per cent in every case.

D. DISCUSSION

The results confirm the earlier observations of Bartley and Nelson (4) and Nelson and Bartley (18) that: (*a*) brightness enhancement occurs when part-spectrum targets are used; (*b*) brightness enhancement occurs more frequently and is of greater magnitude with some wavelengths than others; and (*c*) part-spectrum brightness enhancement is critically dependent upon timing, i.e., rate and cyclic properties of intermittency. On the other hand, the relation of hue-saturation shifts or the washout effect to brightness remains to be measured. It appears as though the brightness peaks reported may not correspond perfectly with the peak washout effect induced by the same portions of the visible spectrum (4).

Because stimulation was parafoveal the sensory end results are associated with two kinds of receptors and a variety of post-synaptic systems. This makes it difficult to relate the effects to any single neuroretinal mechanism.

Nevertheless it might be pointed out that in general, explanation of these results and those previously reported, using foveal stimulation (4, 18, 21), seems to require consideration of mechanisms governing synchronization and amplitude of activity induced by afferent stimulation within the post-retinal portions of the optic pathway. The repeated occurrence of part-spectrum brightness enhancement itself suggests the operation of optic pathway mechanisms utilized by Bartley to explain full-spectrum brightness enhancement, i.e., his "alternation of response" theory (2, 3, 5). Another matter lending weight to this conclusion is the rather close correspondence between specific conditions of PCF and rate effective for both part- and full-spectrum brightness enhancement (18).

The relations between optic pathway activity and the class of sensory end results reported upon here seem potentially of major importance. There are reasons, for example, to expect explanation of part-spectrum results to clarify the nature of brightness enhancement itself. Full-spectrum stimulation falling within the range of intensities used would not have produced brightness indexes so high as were reported here. This probably stems from the fact that all wavelengths do not produce equal amounts of brightness enhancement.

Explanation of part-spectrum brightness enhancement may also provide a theoretical account of the fact that full-spectrum stimulation can produce chroma under proper conditions of intermittency, i.e., the "Prevost-Fechner-Benham subjective color effect" (10). The correspondence between the temporal conditions effective for chromatic brightness enhancement, the Brücke-Bartley effect (full-spectrum brightness enhancement) and Prevost-Fechner-Benham subjective color effects is too close to ignore. Earlier workers (1, 6, 7, 10, 11, 12, 22, 24, 25) report chroma induced with rate of stimulus intermittency between two and 16 cps. Five to 10 cps were commonly reported optimal. Earlier workers also report PCF to be a critical variable. Hurst (14), Burch (8), and Stewart (26) specifically recommend a ratio of .25. Bagley (1) made a systematic investigation of the cyclic variable and concluded "the most striking change is with a disc 3/4 black."

The senior authors have stressed here and elsewhere how part- and full-spectrum brightness enhancement seem to depend upon subfusional rates of intermittency and upon a type of cycling wherein the period of null stimulation is several times greater than that of the photic, i.e., PCF .25.

E. SUMMARY

Brightness enhancement was measurably present for part-spectrum impingements presented parafoveally under the proper conditions of intermittency. Some wavelengths were more effective than others. Full-spectrum brightness enhancement or the "Brücke-Bartley effect," the "Prevost-Fechner-Benham subjective color effect," part-spectrum brightness enhancement, and the "washout effect" are so closely related on an operational level that they seem likely to be products of a common biological mechanism. No attempt was made to discuss the specific biological mechanism responsible although the phenomena are likely related to factors Bartley outlines in his "alternation of response theory."

REFERENCES

1. BAGLEY, F. W. An investigation of Fechner's colors. *Amer. J. Psychol.*, 1902, **13**, 488-525.
2. BARTLEY, S. H. Some facts and concepts regarding the neurophysiology of the optic pathway. *A.M.A. Arch. Ophthalmol.*, II, 1958, **60**, 775-791.
3. ———. A clarification of some of the procedures and concepts involved in dealing with the optic pathway. *In* Jung, R., and Kornhuber, H. (*Eds.*), *The Visual System: Neurophysiology and Psychophysics*. Heidelberg: Springer-Verlag, 1961.
4. BARTLEY, S. H., & NELSON, T. M. Certain chromatic and brightness changes associated with rate of intermittency of photo stimulation. *J. of Psychol.*, 1960, **50**, 323-332.
5. ———. Some relations between sensory end-result and neural activity within the optic pathway. *J. of Psychol.*, in press.
6. BAUMANN, C. Beiträge zur Physiologie des Sehens: IV. Subjektive Farbenerscheinungen. *Pflüg. Arch. ges. Physiol.*, 1912, **146**, 543-552.
7. BRÜCKE, E. Über die Nutzeffect intermitterender Netzhautreizungen. *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften*, 1894, **49**, 128-153.
8. BURCH, G. J. An account of certain phenomena of colour vision with intermittent light. *J. Physiol.*, 1897, **21**, 426-434.
9. COHEN, J., & GORDON, D. A. The Prevost-Fechner-Benham subjective colors. *Psychol. Bull.*, 1949, **46**, 97-136.
10. DONISELLI, C. Sui fenemeni d'induzione cromatic da luce bianca e sulla natura dei processi consecutivi. *Arch. Fisiol.*, 1907, **4**, 561-593.
11. FRY, G. A. Color phenomena from adjacent retinal areas for different temporal patterns of intermittent white light. *Amer. J. Psychol.*, 1933, **45**, 714-721.
12. ———. Color sensations produced by intermittent spectral stimuli. *Amer. J. Psychol.*, 1936, **48**, 326-330.
13. HELMHOLTZ, H. v. Treatise on Physiological Optics. Vol. II. 3rd ed. (*Ed.* J. P. C. Southall). Menasha, Wisc.: George Banta Publishing Co. for the Optical Society of America, 1924, 255-261.
14. HURST, C. H. A spectrum top and a complementary one. *Nature, Lond.*, 1895, **51**, 510.
15. IVES, H. E. Studies in the photometry of lights of different colors. I. *Phil. Mag.*, 1912, **24**, 149-188.
16. ———. Studies in the photometry of lights of different colors. III. *Phil. Mag.*, 1912, **24**, 744-751.
17. ———. Studies in the photometry of lights. IV. *Phil. Mag.*, 1912, **24**, 843-853.
18. NELSON, T. M., & BARTLEY, S. H. The role of PCF in temporal manipulations of color. *J. of Psychol.*, 1961, **52**, 457-477.

19. NELSON, T. M., BARTLEY, S. H., & DEHARDT, D. A comparison of variability of three sorts of observers in a sensory experiment. *J. of Psychol.*, 1960, **49**, 3-11.

20. ———. Considerations involved in obtaining averages of CFF data. *J. of Psychol.*, 1960, **49**, 263-277.

21. NELSON, T. M., BARTLEY, S. H., & MACKAVEY, W. R. Responses to certain pseudo-isochromatic charts viewed in intermittent illuminance. *Percept. & Mot. Skills*, 1961, **13**, 227-231.

22. PIÉRON, H. Le méchanisme d'apparition des coleurs subjectives de Fechner-Benham. *Annee Psychol.*, 1922, **23**, 1-49.

23. PURDY, D. M. Spectral hue as a function of intensity. *Amer. J. Psychol.*, 1931, **43**, 541-559.

24. ROOD, O. N. Subjective optical phenomena. *Amer. J. Sci. & Arts*, 1861, **31**, 417.

25. STEWART, G. N. Is the law of Talbot true for very rapidly intermittent light? *Proc. Roy. Soc. Edinb.*, 1887-88, **15**, 441-464.

26. ———. Color phenomena caused by intermittent stimulation with white light. *Amer. J. Physiol.*, 1924, **69**, 337-353.

Department of Psychology
Michigan State University
East Lansing, Michigan

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 397-400.

DAILY PREFERENCES AND INSTITUTIONALIZATION*[1]

Veterans Administration Center, Mountain Home, Tennessee

HARRY F. DESROCHES AND CHARLES L. WALTER

A. INTRODUCTION

Farber (1) presented an interesting hypothesis that it is the expectation of pleasantness or unpleasantness of future events which determines preferences for particular days of the week. He had 80 college students rank the days of the week in order of preference. Farber's students were divided into two groups: those with Saturday classes and those without. For both groups, Monday was the least preferred day with an orderly increase in preference as Saturday approached. Both groups ranked Saturday as the most preferred day. The groups differed in their ranking of Friday and Sunday. The students without Saturday classes ranked Friday as the second most preferred day and Sunday as third most preferred day. The students with Saturday classes reversed their ranking, giving the second place to Sunday and the third to Friday. Farber interpreted the results in terms of expectations. Thus, Saturday itself was most preferred because both groups had the expectation that the next day, Sunday, would be a leisure day. Sunday was less preferred than Saturday because of the expectation that the following day would begin the work week. Since Friday, a work day for both groups, was ranked so high in preference, leisure on the day itself was not the main factor generating differential preferences. Farber interpreted his results as supporting his hypothesis relating expectation to daily preference, and, more generally, as supporting a contention that one's feeling tone is determined more by future than by present activities.

Differential ranking or differential preferences for days of the week presupposes differential stimulation and differential activities during the week. For some members in an institution differential stimulation and differential activities are not dependent on any particular day of the week. One of the big changes for these people on coming into an institution is a loss of the

---

* Received in the Editorial Office on December 16, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors wish to express their appreciation to Drs. E. R. Larsen, B. D. Kaiman, Miss Sarah Booher, and Mr. Werner Berg for their helpful comments on the text, to Mr. Earl Thompson for his statistical calculation, and to the veterans and staff at Mountain Home Domiciliary for their cooperation and interest.

weekly structure imposed outside the institution by work activities. For those members who are too disabled to move about in their environment, habituation to environmental cues produces a monotonous sameness, which cannot be overridden by small changes due to different daily stimulation. Thus, if a group of people do not have different expectancies of pleasantness for different days of the week because they lack differentiating stimulation and activities for the days of the week, no consistent daily preferences should be found for this group of people.

For other members in an institution who consistently do different things on different days of the week and who are physically capable of moving freely in their environment, consistent daily preferences should emerge. The daily preferences should be more consistent in a group of members who work an eight-hour day, five-day week, and get a small pay for such work (member employees) than in a group of members who work a six-hour detail, five days per week for no pay (detail employees) since the member employee group receives more differential stimulation.

Following from this theoretical analysis, it was predicted that in a VA domiciliary subjects whose activities are so restricted they have no work assignments (NW) will not show group trends or consistency in daily preferences, while subjects who do have work assignments will show consistent group preferences for particular days of the week. Within the latter group it was predicted that member employees (ME) will show more consistent daily preferences than detail employees (DE). It also was predicted that the ranking of the days of the week by those subjects who have work assignments would be similar to the rankings obtained by Farber.

B. PROCEDURE

Three samples were drawn from men institutionalized at the VA Domiciliary at Mountain Home. The first was a random sample of the most disabled veterans. These disabled subjects (NW) have no detail or work assignments and most of them do not stray far from their buildings. The second and third samples were taken from a special ward. The men on the special ward are considered to be the best rehabilitative candidates in the domiciliary. They work six hours per day on special details (DE, sample 2) or eight hours per day for a small pay check (ME, sample 3). There were 25 subjects in the first sample, 14 subjects in the second sample, and 11 subjects in the third sample.

Subjects were told that they were participating in an experiment and that the experimenter was interested in which day of the week they liked best and which days of the week they liked least. There were two stages to obtaining

TABLE 1
MEAN GROUP RANKINGS FOR THE DAYS OF THE WEEK BY COLLEGE SUBJECTS AND BY INSTITUTIONALIZED MEN

	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
College students							
With Saturday classes	3.0	5.9	5.0	5.0	4.4	3.2	1.5
Without Saturday classes	3.1	6.3	5.0	4.8	4.3	2.8	1.6
Institutionalized Men							
Member employees	2.6	5.4	5.1	4.9	4.5	3.3	2.2
Detail employees	2.9	5.0	4.2	4.4	4.5	4.4	2.7
Non workers	2.9	3.8	4.1	4.3	4.2	4.7	4.0

the data. In the first stage, subjects were told to rank the days of the week by placing the number "1" beside the day most preferred, the second rank to the day next preferred, etc. If these instructions did not yield a ranking, the second stage was attempted. In the second stage, subjects were asked which day they preferred most, followed by which day they preferred least and then the second most preferred day and so forth. For those days which the subject could not rank, each day was given the average rank of the remaining ranks.

C. RESULTS AND DISCUSSION

Table 1 gives the mean group rankings for the various days of the week by Farber's two groups of college students, and by the three samples of institutionalized men at Mountain Home Domiciliary. An analysis of variance for repeated measurements yielded a significant F for the ME group ($F_{6,\,70} = 8.22$, $p < .01$), a less significant F for the DE group ($F_{6,\,91} = 2.41$, $p < .05$), and nonsignificance for the NW group ($F_{6,\,168} = 1.31$, $p > .05$). Thus, the experimental hypotheses were supported. The ME group showed high consistency in their daily preferences, the DE group lesser consistency in their daily preferences, and the NW group showed no significant consistency.

The mean rankings of the days of the week by the Domiciliary groups were correlated (Spearman Rank Correlation) with the mean rankings by the group of college subjects without Saturday classes yielding the following rho's: ME $+.96$ ($p < .01$), DE $+.63$ ($p < .10$), and NW $-.14$ (ns). Thus, rankings by the working samples at Mountain Home resembled in the way predicted the mean rankings of the days of the week by college subjects.

Both Farber's study and the present study indicate that daily preferences are not due to chance, but rather are due to the differential stimulation and/or activities of the various days of the week. The weekly structure provided by a school week or a work week, probably because of expectancies of either pleasant and/or unpleasant events, produces differential preferences for days of the week. The monotony of inactivity and confinement and its expression as "one day is like the other" has often been described in fiction and personal accounts. It is surprising, nevertheless, how sensitive an index daily preference is to institutionalization.

REFERENCE

1. FARBER, M. L. Time perspective and feeling tone: A study in the perception of days. *J. of Psychol.*, 1953, **35**, 253-257.

VA Center
Mountain Home
Tennessee

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 401-408.

COMBINING EXTERNAL AND INTERNAL SYMBOLIZATION IN THE LSD EPISODE*[1]

Arkansas State Hospital

DONALD K. GUCKER

A. PURPOSE

The usefulness of D-lysergic acid diethylamide (LSD-25) with respect to facilitation of psychotherapeutic endeavor has been adequately established (1, 4, 5, 6, 7, 10, 12, 14, 16, 17, 20, 21, 23, 24, 26, 27, 28, 30). In particular, therapists stress the unique value in relation to emergence of preconscious and unconscious material, ego enhancement, dramatization of conflictual material, acceleration of treatment, effect of transference intensity, disruption of the defensive system, retention of the intact ego during treatment, and intensive abreaction (2, 3, 4, 11, 21, 22, 25, 26). Perhaps the most significant impression is that patients not found amenable to the usual therapeutic approaches will at times respond to treatment with LSD (8, 10, 23, 27). Needless to say, the particular value is dependent on the particular patient. With certain nosologic categories, the use of LSD is contraindicated, and even within those categories that have been found amenable to treatment with LSD, the potential value is dependent on the particular personality structure (13, 28).

Internal symbolization (IS) refers to the hallucinatory material produced by the patient while under the influence of LSD, and external symbolization (ES) refers to objects or situations, interjected by the therapist into the patient's perceptual field during the drug interview. Much mention of hallucinatory material is found in the literature, but little mention is made of the specific significance and potential therapeutic utility of IS. Eisner and Cohen (17), in recognizing these factors, state, "One of the more therapeutic actions was to direct the patient toward the frightening or painful symbolic material and to assist him in experiencing and understanding it." The similarity between the hallucinatory material produced in the LSD episode and dream symbolization has been mentioned by Sandison (23).

---

* Received in the Editorial Office on December 17, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The content of this paper is based on personal contact with approximately a thousand LSD episodes, conducted in a private psychiatric setting.

Internal symbolization seems to be a product of the ego's defensive organization, and just as in the dream, unconscious material finds expression in a symbolic or resistive form. A therapeutic advantage offered by LSD is that the process of defensive symbolization takes place in the presence of the patient's *intact* ego. Level of dosage, frequency of treatment, and concomitant psychotherapy may be manipulated so the ego will remain relatively intact during the drug episode. On rare occasion, a patient would be forced into a dissociative defensive maneuver (1 to 24 hours in the author's experience) but no lasting negative effect was apparent. Savage (29) has published a point of view somewhat at variance with this conclusion. Cohen (13) has made available the experience of many therapists with respect to the risk involved in the use of LSD.

The nature and use of ES lacks general recognition in the literature concerned with LSD. Eisner and Cohen (17) cite the use of music, mirrors and family photographs in their LSD treatments. External symbolization has long been used in the form of play therapy with children, and subtle ES is often found in adult therapy. In this work, many symbols were lifted from the play therapy situation and applied to the LSD treatment of adults. The baby bottle, clay, dolls, art media, use of the toilet, photographs, and TAT cards were employed. External symbolization of a situational nature was frequently introduced in the form of social interactions: between therapist and spouse in the case of marriage partners, prearranged situations between therapists, and the use of transference relationships that had previously been set up between the patient and other hospital personnel. Frequently the ES was absorbed into the IS, but at times the patient would not permit this to occur.

Resistance in the form of IS may be useful if appropriate ES in introduced, and if interaction between IS and ES occurs. Symbolic resistance may thus be seduced and used as a tool for therapeutic growth.

The combination of IS and ES can provide a bridge between discharge of libidinal material in hallucinatory form and the difficult working through of conflictual material in the external world. Such a philosophy requires the therapist to assume a directive function, especially with respect to selection and introduction of the external symbols and interpretation of the IS. The choice of ES may be aided by the content of the patient's IS but generally, knowledge of the basic conflicts as reflected in history, dreams and transference phenomena are an adequate guide.

Abreaction may occur with respect to the ES; for example, hostility may be ventilated toward a doll, or directly toward the therapist. He may be involved as both IS and ES, depending upon transference needs and the role he chooses. A baby bottle may be perceived as a source of oral satisfaction, a symbol of mother, a focus for hostility, an object for self-nourishment or whatever the patient needs it to be. The mere sight of the toilet may trigger feelings of hallucinatory material concerned with this particular phase of development.

B. PROTOCOL

The following transcript illustrates a combination of ES and IS with respect to an attempt toward emotional self-nourishment. This adult female patient was troubled by an alcoholic problem, depression, suicidal ideation, and in general deep hostility turned both toward herself and others. Long term psychoanalytically oriented psychotherapy with LSD as an adjunct had preceded the present episode, but with modest gain. Sexual misidentification is evident, and the expression and recognition of this problem in this session is interesting. As usual the transcript was reviewed with the patient the day following the LSD episode.

Dosage—200 micrograms—orally—LSD at 1:15 p.m.

Patient speaks of not liking the room she is in today—doesn't like the bed or the way the room is arranged. She thinks the medication may not work because she ate only an hour before. "My ears feel funny, like they are about to pop. Maybe I'm going on a space flight." (She laughs.) "I'm awfully scared of dying. Sometimes I get so scared I'm almost sick. Sometimes—I think I'm going to die soon." (She laughs again.) "I've been thinking about it a lot the last three days. It seems that if I get well I'll have to die. Sometimes, especially in the evening when I am walking outside and I see the sunset, I think, 'My God, how terrible to die!' It scares me so I get sick to my stomach. I imagine myself being killed in an automobile wreck. I imagine myself being all horribly mutilated and dead." (She laughs.) "I cannot imagine myself hurt a little bit. I always have to be dead. I feel real hungry now like I was very, very little, and I am lying in a bed and it is cheerful, and I am getting well from something. That's a funny thing to imagine at the same time as imagining yourself dead. I much prefer getting well, but it seems I have to die whether I want to or not. Sometimes I imagine what it would be like to be killed, if there was another nuclear war. I see my kids and myself dead. Then I get mad at myself for wasting my time on such horrible thoughts. I think I'd go nuts if anything happened to the kids. I don't think I could take it if anything happened to them." (Patient yawns.) "Gee, I'm getting a headache. My ears are ringing." (Patient holding hands clasped in prayer position.) "I can't figure out what's going on. My head is spinning." (She smiles and sits up on the side of the bed.) "Do you suppose they'll give me a couple of aspirin? I don't like the way this bed sinks in the middle. You can't swing your legs." (Patient laughs.) "It's just not built for it. Besides, it's facing the wrong direction. God, I feel funny!" (Patient drops her eyes.) "Wish I could find out what makes me have such bad dreams in the morning. Sometimes I feel the greatest test as whether I am well or not will be whether I have bad dreams in the morning. God, I have a splitting headache! I guess I'm afraid of something—though I'd hate the thought of going out of this place being still afraid of something. I thought this was all done with. I don't want sympathy. I'd much rather be well and able to operate by myself. My skin feels like jelly, almost feels like I haven't got any spine." (Patient laughs.) "It's like I am being dissolved." (Laughs again.) "I feel helpless with this hand in a splint, have to do everything with my left." (Squeezing her eyes tightly shut.) "Just trying to figure out what's happening." (Lies back against the wall, covers her legs with a blanket.) "My teeth are chattering. Everything in this room seems the wrong way. Everything is on the wrong side of me." (The room is actually arranged the opposite of her usual room.) (Patient curls up on her side with her eyes closed.) "Are you supposed to say whatever you think about in these things?" (Therapist in. Patient tells him she isn't getting anywhere.) "I keep wondering what it would be like to have a baby, just for myself, maybe, another immaculate conception." (Patient laughs.) "I don't know if I feel old or young. I'm glad you're here today. Wish you could help me find out what the monsters are. I don't feel like it has to do with anybody. I feel very much alone. Afraid of nothing in particular." (Therapist out.) "I felt safer with him in here than I do now. It's sort of like being afraid to have a baby. If I ever had myself, I would take care of it, take better care of myself, like shaving my legs tonight." (Patient laughs.) "I keep thinking about a little baby and how I will take care of it. I seem to feel myself big and small, like two separate people, like the big one has to take care of the little one—somehow. But the big one is so scared, it doesn't know what to do. Just the sight of that little baby is a disturbing thing. It doesn't even have a sex. I have to give it one. I have to decide whether it will be a man or a woman. I'd almost rather it was a man." (Patient lying with her head at the foot of the bed.) "I still wish I had been a boy—like my father. Seems like I have to give myself my own sex—whether I like it or not. It's just the baby's big open mouth. I can't tell it to stop. I'd like to go back looking like a baby. But it looks like a cartoon with a big, open mouth. It's like a cavern that you can't fill. It's like you couldn't stick such a little thing as a bottle in that big mouth. You'd have to stick a penis in it or something. But if I'm not a baby, I don't want that either. I suppose you feed them or slap them—to get the baby to close its mouth. If I can't make this mouth go back to a normal mouth, I'll never be able to feed it. I'll have to make it into a nice little girl baby. It's still got a big mouth." (Patient laughs.) "If you put a bottle in it, it would just disappear. Maybe I should make an extra big bottle for it." (Patient laughs.) "A giant size bottle. The bigger the bottle gets, the bigger the mouth gets." (Patient laughs and yawns.) "I don't know why I've been thinking of a bottle. It could have a great big breast.[2] In all the world, you should be able to find one big enough. If I don't figure out what to do, that baby is going to die. It's almost like I have to make a smaller one so it won't want so much." (Patient laughs.) "How do you give a baby a bottle anyway?—Guess you just stick it in its mouth. I've forgotten what milk tastes like. I feel like there isn't a breast or penis in the whole world that could fill it. Besides, it wants something to eat, not to suck on. It wants a lot of milk." (Patient wants some milk. Nurse brings milk to her. Patient says it tastes good and wants more.) "It just doesn't want milk. It wants to suck, too. Guess I have to drink it from a glass. It wants to suck so I have to hate it. Maybe it's not so silly. Maybe it's the only logical thing to do." (Patient wants a bottle, drinks another glassful of milk while she's waiting for the bottle.) "I don't think it matters if it is a girl. I'll love it just the same, even maybe 'cause it's mine." (Patient lies down and covers herself, sucking from a baby bottle. She didn't want the nurse to give it to her, said she'd give it to herself.) "This feels sort of silly, but it's nice. I like it. It's nice that I don't want anyone to do it for me. I want to do it myself. I'm glad the baby is a girl. It's nice having a girl and not being ashamed of it. When I tried it the other night I didn't know how to suck on it, but now I do." (Patient laughs.) "There's a knack to it so you don't get your teeth in the way." (Patient laughs again.) "I am getting full." (She is finishing her third carton.) "I feel good. It seems so wonderful and so simple and so nice to do it for myself. Maybe this room wasn't so bad after all. I feel too good to even go to sleep." (Patient sits up on the side of the bed to light a cigarette, sees that she left a small amount of milk in the bottle, and lies down to finish it.) "What time it it? I'm not afraid any more. Oh, I guess I am, sort of.

---

[2] The equation of penis and breast is interesting considering Freud's remarks on this subject (19).

I really enjoyed sucking on that bottle. Nobody told me I shouldn't or that I wasn't supposed to. The baby doesn't have his mouth open any more. I'm just not used to calling myself 'she'." (The patient goes to the bathroom.) "This is kind of a nice little room after all. I feel so nice and warm and peaceful. This room is sort of special. I'll have to bring my kids here some day and show them where I was born." (The patient laughs.) "I feel like saying, 'That wasn't so bad, was it?' In fact, it's kind of nice being born. I just thought I saw a spider run across there, and it's nothing but a work in the wood. Even if that would have been a spider, I wouldn't have let it go. It wouldn't hurt me. It is only 4.15. It's funny what you do to yourself, only I think I started out right with this baby this afternoon. I think she should be very happy that I found her. If I don't sleep well tonight, I will just punch them in the nose." (She burps.) "There! I burped." (Patient laughs.) "It seems so simple now. If somebody's mouth is open, just feed it. That's what they kept telling me, but I didn't believe that. It seems funny to sit here and look at my unshaven legs and think, 'My God, I'm a woman'!" (Patient laughs.) "I have to wipe my own butt and powder it, but it doesn't seem so bad. It's not always so simple, but it's good enough for me at the moment. I feel like celebrating. I'm glad my therapist wasn't here. I'm glad I did it by myself. I really feel like going to sleep except I am too excited. I feel like saying, 'Look what I did!' yet, it's not so much after all."

4:45 p.m.—Termination medication given. (100 mg. Thorazine)

C. CONCLUSION

The transcript illustrates the integration or working through of conflictual material in a way that facilitates ego development. Symbolic expression of need is dealt with by introducing ES into the fantasy created by the interaction of the primary and secondary processes and the required regressed state is achieved through the use of LSD. This is not to say that most therapeutic endeavor does not employ symbolic interaction, but that such therapy does so without the intensity, directness, and speed available with LSD. Again it must be pointed out that this is not so in all cases. The literature, at times, has been a bit grandiose in its claim, but adequate guidelines and prohibitions concerned with the selection of patients are available in the aforementioned references.

REFERENCES

1. ABRAMSON, H. A. Lysergic acid diethylamide (LSD-25): III. As an adjunct to psychotherapy with elimination of fear of homosexuality. *J. of Psychol.*, 1955, **39**, 127-155.

2. ———. Lysergic acid diethylamide (LSD-25): XIX. As an adjunct of brief psychotherapy, with special reference to ego enhancement. *J. of Psychol.*, 1956, **41**, 199-299.

3. ———. Lysergic acid diethylamide (LSD-25): XXII. Effect on transference. *J. of Psychol.*, 1956, **42**, 51-98.

4. ———. Verbatim recording and transference studies with lysergic acid diethylamide (LSD-25). *J. Nerv. & Ment. Dis.*, 1957, **125**, 444-450.

5. Abramson, H. A. (*Ed.*), The Use of LSD in Psychotherapy. New York: Josiah Macy, 1960.

6. Arendsen-Hein, G. W. Hallucinogenic drugs. *Lancet*, 1961, **1**, 445.

7. Busch, A. K., & Johnson, W. C. LSD-25 as an aid in psychotherapy. *Dis. Nerv. System*, 1950, **11**, 241-253.

8. Cameron, K. Hallucinogenic drugs. Specific problems. *Lancet*, 1961, **1**, 445.

9. Cattell, J. P. Use of drugs in psychodynamic investigations. *In*: *Experimental Psychopathology.* New York: Grune & Stratton, 1957, p. 218.

10. Chandler, A. L., & Hartman, M. A. Lysergic acid diethylamide (LSD-25) as a facilitating agent in psychotherapy. *Arch. Gen. Psychiat.*, 1960, **2**, 286.

11. Cohen, S. Notes on the LSD state. *J. Wadsworth Gen. Hosp.*, Los Angeles, 1959, **3**, 79.

12. ———. The therapeutic potential of LSD-25. *In*: *A Pharmacological Approach To The Mind.* Springfield, Ill.: Chas. C Thomas, 1959, p. 251.

13. ———. Lysergic acid diethylamide: Side effects and complications. *J. Nerv. & Ment. Dis.*, 1960, **130**, 30.

14. Cohen, S., & Eisner, B. G. Use of lysergic acid diethylamide in a psychotherapeutic setting. *Arch. Neurol. & Psychiat.*, 1959, **81**, 615.

15. Cohen, S., Fichman, L., & Eisner, B. G. Subjective reports of lysergic acid experiences in a context of psychological test performance. *Am. J. Psychiat.*, 1958, **115**, 30.

16. Cutner, M. Analytic work with LSD-25. *Psychiat. Quart.*, 1959, **33**, 715.

17. Eisner, B. G., & Cohen, S. Psychotherapy with lysergic acid diethylamide. *J. Nerv. & Ment. Dis.*, 1958, **127**, 528.

18. Field, M., Goodman, J. R., & Guido, J. A. Clinical and laboratory observations on LSD-25. *J. Nerv. & Ment. Dis.*, 1958, **126**, 176.

19. Freud, S. Collected Papers. London: The Hogarth Press, 1949, **3**, 64-65.

20. Langner, F. W., & Kemp, M. D. *Quart. Bull., Pinebluff San.*, 1956, **1**.

21. Ling, T. M., & Buckman, J. The use of lysergic acid in individual psychotherapy. *Proc. Roy. Soc. Med.*, 1960, **53**, 927.

22. Polo, A., Krinsky, L. W., & Goldbarb, L. LSD as an adjunct to psychotherapy with alcoholics. *J. of Psychol.*, 1960, **50**, 85.

23. Sandison, R. A. Psychological aspects of the LSD treatment of the neurosis. *J. Ment. Sci.*, 1954, **100**, 508.

24. ———. The clinical uses of lysergic acid diethylamide. *In*: *Lysergic Acid Diethylamide and Mescaline In Experimental Psychiatry.* New York: Grune Stratton, 1956, 27.

25. ———. Comparison of drug induced and endogenous psychosis in man. *In*: *Neuro-Psychopharmacology.* New York: Elsevier, 1959.

26. ———. Hallucinogenic drugs. Therapeutic applications. *Lancit*, 1961, **1**, 444.

27. Sandison, R. A., Spencer, A. M., & Whitelaw, J. D. A. The therapeutic value of lysergic acid diethylamide in mental illness. *J. Ment. Sci.*, 1954, **100**, 491.

28. Sandison, R. A., & Whitelaw, J. D. A. Further studies in the therapeutic value of lysergic acid diethylamide in mental illness. *J. Ment. Sci.*, 1957, **103**, 332.

29. SAVAGE, C. The resolution and subsequent remobilization of resistance by LSD in psychotherapy. *J. Nerv. & Ment. Dis.*, 1957, **125**, 434.

30. WHITELAW, J. D. A. A case of fetishism treated with lysergic acid diethylamide. *J. Nerv. & Ment. Dis.*, 1959, **129**, 573.

Arkansas State Hospital
4313 West Markham Street
Little Rock, Arkansas

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 409-418.

THE RETENTION OF INCIDENTALLY LEARNED CONNOTATIVE RESPONSES*[1]

Norwich Hospital

HALIDE S. YAVUZ

A. INTRODUCTION

Clinical evidence as well as everyday experience indicate that even though an often emitted name of a person or the verbal label of an object cannot be recalled, it may still be possible to make a valid evaluative statement regarding the person or the object. Thus one may justifiably say, "I do not recall his name, but he was a good man." If it can be assumed that the supposed maker of the foregoing statement never had a set or intention to learn an evaluative response to the person in question, it would follow that the emission of the word *good* was contingent on some sort of prior incidental learning. It is here proposed that explanations of such effects can be deduced from current verbal learning theories. That the phenomenon can be made experimentally manageable was shown in an earlier study by the author and Bousfield (7). In keeping with this point of view, the present study was undertaken as a test of the following experimental hypothesis: Subsequent to the learning of a given meaningful response presented as a translation of a paralog, there should be a tendency for the connotatively meaningful evaluative responses acquired by the paralog as a result of this learning to be retained when the verbal response itself can no longer be recalled.

Actually two separate experiments were undertaken and both yielded essentially similar data. The present account describes the first only. The procedure involved having subjects learn paired associates, such as, ASLAN-SUNLIGHT. The stimulus term of this pair is a Turkish word and the response was presented as its English translation. The prediction was that in the event the subject forgot the purported meaning of ASLAN, he should show, nevertheless, a tendency to retain an evaluative response acquired by ASLAN during the initial learning. The predicted tendency was measured

---

* Received in the Editorial Office on December 17, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This paper is based upon a doctoral dissertation submitted to the Graduate School of the University of Connecticut in 1961. The author would like to express her sincerest and deepest appreciation to her major adviser, Weston A. Bousfield, and to her other advisers, James M. Sakoda and Sam L. Witryol, who, with their critical comments and encouragements, guided the progress of this dissertation.

by allowing time for forgetting and then having the subject rate ASLAN on the good-bad semantic differential scale. It was assumed the hypothesis would be supported when forgetting occurred and the subject conformed to the normative ratings of SUNLIGHT by rating ASLAN toward the good end of the scale.

Two current theories of verbal behavior appear to provide adequate explanations for the phenomenon here considered. One is Osgood's (5) mediational theory. He regards the connotative meaning of a word as some sort of behavioral pattern elicited originally by an unconditioned stimulus and transferred to the given word by conditioning. The behavioral pattern is elicited immediately by the given word and it produces stimulation to which instrumental and conditioned responses can become attached. In these terms the subject rates SUNLIGHT as having a good meaning on the basis of some sort of judgment he makes of the quality of the behavioral state which SUNLIGHT elicits. As applied to the learning of the paired associates ASLAN-SUNLIGHT, the primary established connection should be between ASLAN and the behavioral pattern representing the connotative meaning of SUNLIGHT. It follows from these considerations that when the subject has forgotten SUNLIGHT as the learned verbal response to ASLAN, he may show, nevertheless, some degree of retention of the connotative meaning which ASLAN should have acquired in the original paired associates learning.

A second set of formulations for explaining the experimental hypothesis may be found in Bousfield's (1) partial response identities theory. According to this interpretation, the word SUNLIGHT in the foregoing example should elicit a composite of implicit verbal associational responses, e.g., moonlight, shine, bright, tan, beach, pleasant, and good. It may be noted that these responses include: (*a*) the word *good*; (*b*) pleasant, which is equivalent to *good* as a response; and (*c*) several words which should be capable of mediating *good* as a response. When the subject learns the pair, ASLAN-SUNLIGHT, the nature of the learning is such as to develop connections not only between ASLAN and SUNLIGHT, but also between ASLAN and the composite of associates elicited by SUNLIGHT. It would follow that when the subject cannot recall SUNLIGHT as the response he learned to ASLAN, he may, nevertheless, be able to rate ASLAN as *good* on the semantic differential scale. He need only recognize *good* as a response having an associative relationship to ASLAN. This rating does not depend on the retention of any single associated response since it presumably can be mediated by several alternative associates, e.g., bright and pleasant. Because the semantic differential actually makes use of the method of recognition, it is highly sensitive in detecting the presence of weak habits. These factors indicate that very little retention is required for the rating of ASLAN appropriately on the semantic differential scale. While the two theories here summarized both appear to be in accord with the experimental hypothesis, it should be noted that the present study provides no basis for judging their comparative merits.

B. METHOD

1. *Subjects*

The experimental subjects were 25 undergraduate and graduate students at the University of Connecticut. Two additional groups comprising undergraduate subjects supplied normative data. A total of 40 subjects provided semantic differential ratings, and 150 subjects provided ratings of meaningfulness.

2. *Stimulus Items and Apparatus*

The stimulus items comprised 18 Turkish words and 18 English words. The former, here designed as paralogs, were as follows: ASLAN, BARDAK, BAYRAM, DEMET, FILIZ, ILIM, HASRET, KALEM, LALE, MERKEZ, NAKIL, ORTAK, SANAT, TEZAT, VATAN, YALAN, ZARIF, ZAHMET.

The English words were chosen from the semantic differential norms of Jenkins, Russell, and Suci (2). These words were divided into three groups of six each and were selected so as to represent respectively three levels of evaluative meaning, namely, Good, Neutral and Bad as shown in Table 1 where they appear along with their Thorndike-Lorge (6) *L* (magazine) and *S* (semantic) counts.

The paralogs and the words were typed on 3 × 5 inch cards for paired associates learning. Each card contained one pair with a paralog as the stimulus item on the left and a word as the response on the right. A different randomized assignment of the words to the paralogs was prepared in advance for each subject. A 3 × 3 foot screen of plywood with a 1 × 2 inch aperture in the center was used for exposing the cards to subjects. A manually operated shutter was set to open in two stages. The first stage permitted exposure of the paralog only, and the second exposed both the paralog and the word with which it was paired. The time intervals for exposures were approximately two seconds for the stimulus item, two seconds for both items. The interval between each card was about one second.

TABLE 1
STIMULUS WORDS WITH THEIR CORRESPONDING SEMANTIC (*SD*) VALUES AND THORNDIKE-LORGE *L* (MAGAZINE) AND *S* (SEMANTIC) COUNTS

		Good				Neutral				Bad	
Word	*SD*	*L*	*S*	Word	*SD*	*L*	*S*	Word	*SD*	*L*	*S*
Faith	1.40	459	412	Dim	4.20	156	168	Devil	6.63	225	211
Joy	1.60	418	423	Heavy	4.10	680	966	Hate	6.53	756	234
Justice	1.67	181	564	Slow	4.00	434	321	Hungry	5.17	274	153
Nice	1.50	830	176	Stout	4.00	99	141	Mad	5.90	400	143
Success	1.50	573	625	Winter	4.10	610	576	Trouble	6.10	1180	734
Sunlight	1.43	160	103	Knife	4.30	251	122	Steal	6.67	169	86
Means	1.51	436.8	383.8		4.11	371.7	382.3		6.16	500.7	260.2

3. *Procedure*

The experiment was conducted in two phases, with the first for learning, and the second for testing. In the learning phase each subject received standard instructions for paired associates learning. He was told, however, that his task was to learn the meanings of a group of foreign words. A practice list of four pairs of items, not used in the experiment proper, was employed to familiarize subjects with the routine and to insure that the instructions were understood. A correction procedure was used in the learning. After any purported translation had been correctly anticipated three times in succession, the pair of which it was a member was removed from the list. When the last pair had been learned, two additional presentations of all 18 pairs were given.

The testing phase, which took place one week after the learning, comprised two parts. Each subject was first asked to recall the purported translation of each paralog as it was presented in the apparatus. About 10 seconds were allowed for each recall. All responses were recorded as correct, incorrect, or no recall. The subject was then given a small sheet of paper on which the exposed paralog was printed along with the standard seven-step, good-bad semantic differential scale. The rating of the given paralog was obtained regardless of whether the subject was able or unable to recall the word he earlier learned as its translation.

C. RESULTS

The experiment yielded two sets of data, namely, the responses the subjects made when they undertook to recall the purported translations of the paralogs, and the semantic differential ratings subjects made of the paralogs whether they recalled or did not recall the translations. The responses the subject made in the interval allowed for recall were in each instance entered under the connotative category of the translation word with which the given paralog had been paired in the initial learning. These responses were designated as correct or incorrect. Alternatively, the notation, "no recall," was made when the subject made no response in the prescribed time limit.

The first of the several analyses of these data was made in order to determine the extent to which the connotative values of the incorrect recalls tended to match the connotative values of the translation words they displaced. For example, one of the initially learned pairs was HASRET-JOY. If instead of giving JOY as the translation of HASRET, the subject gave the incorrect response MERCY, a matching of connotative value was said to have taken place. Both JOY and MERCY are typically rated by subjects as having a good connotative meaning. A significant tendency for such matches to occur should be regarded as evidence supporting the experimental hypothesis. The data indicated that 21 of the 25 subjects gave a total of 65 incorrect responses. These comprised 34 intrusions of new words, and 31 instances of generalization of responses to the wrong stimuli. In keeping with the plan for the analysis, a group of 40 subjects was asked to give semantic differential ratings on the good-bad scale of all the responses classed as incorrect, i.e., the 31 intrusions. They also rated the 18 original translation words. These responses were then classified on the following basis: Good, if the mean semantic differential rating were between 1.00 to 3.50; Neutral, if the rating were between 3.51 and 4.50; Bad, if the rating were between 4.51 and 7.00. Each incorrect response was further classified as: appropriate, if its classification matched that of the word it displaced in recall; inappropriate, if its classification did not match that of the word it displaced. In these terms, 43 of the incorrect recalls were classed as appropriate and 22 as inappropriate. The mean number of appropriate responses per subject was 2.05, whereas the mean number of inappropriate responses per subject was 1.05. A *t* test of the difference between these means indicated significance at the .05 level. Thus, when subjects responded incorrectly to the paralogs by making errors of substitution, they showed, nevertheless, a significant tendency to substitute words similar in connotative meaning to the words they displaced.

As indicated earlier, subjects were allowed a period of 10 seconds to recall the purported translations of the paralogs. Furthermore, they rated each paralog on the semantic differential scale regardless of how they responded to it. An analysis of the data was undertaken in order to determine the nature of the semantic differential ratings when subjects were unable to give any response. These cases were given the label of no recall. The means of the instances of no recall, classified in terms of the connotative categories of the words with which the paralog had been paired in the initial learning, were as follows: Good, 1.66; Neutral, 4.60; Bad, 6.15. The corresponding means of the semantic differential ratings which the subjects gave to the paralogs in the instances of no recall were: Good, 3.28; Neutral, 4.28; Bad, 4.54. These ratings of the paralogs when subjects were unable to give word associations were necessarily critical from the standpoint of the experimental hypothesis. The use of *t* tests of the differences between the means for no recall gave the following results. The difference between the means for Good and Bad was significant at the .01 level. The difference for the

Good and Neutral categories was significant at the .02 level. The difference for the Neutral and Bad categories, however, was not significant. It is apparent that since low semantic differential ratings are in the direction of *good*ness and high ratings are in the direction of *bad*ness, it may be inferred that the experimental hypothesis receives support from this treatment of the data.

An analysis was undertaken of the instances in which the purported meanings of the paralogs were correctly recalled. As was expected, the ratings of the connotative values of the paralogs were not significantly different from the ratings reported for the purported translation words as obtained from the norms of Jenkins, Russell, and Suci (2). Also, the experimentally obtained ratings did not differ significantly from those given to the translation words by the group of 40 subjects who supplied normative data for the present study. It was found, however, that the mean number of correctly recalled connotatively Bad words, 2.04, was significantly lower ($p < .01$) than the means of the Neutral, 3.04, and the Good, 3.04. The relatively poor recall of the Bad words does not appear attributable to their frequencies-of-usage as assessed on the basis of the Thorndike-Lorge (6) counts. An attempt had been made to select the three connotative categories of words so they would be as equivalent as possible in this respect. No conclusive explanation of the relatively poor showing of the Bad words in recall appears possible. The question did arise, however, of the possibility that the Bad words might differ in meaningfulness or association value from the words in the Neutral and Good categories. Noble's (3) analysis appeared to be especially relevant to this question. He not only devised a simple procedure for measuring meaningfulness but demonstrated that his measure as applied to verbal items correlated significantly with their memorability (4). Noble's measure of meaningfulness of an item is the mean number of acceptable written verbal associative responses it elicits from a group of subjects within a period of one minute. In carrying out this plan, Noble's measure of meaningfulness was obtained for the 18 translation words used in the present experiment from a group of 150 undergraduate student subjects. The following mean values were derived from the data for the three categories of words: Bad, 8.63; Neutral, 10.05; Good, 9.56. The use of t tests of the differences between these means indicated that, while the difference between the Good and Neutral was not significant, the difference between Bad and Good, and between Bad and Neutral was significant at the .05 level. Thus it would appear that the Bad words which as a group were least well recalled had less meaningfulness as defined by association value than the Good and the Neutral words.

D. DISCUSSION

It has been shown in this study that it is possible for subjects to forget verbal responses they have repeatedly made to a given stimulus, and nevertheless still be able to assign an appropriate connotative value to this stimulus. The necessary retention occurs even though no observable evaluative responses to the given stimulus have been emitted. It occurs, moreover, in the absence of specific instructions to learn, and presumably in the absence of a set to learn. The important question for theory is that of explaining the apparently paradoxical fact that under some conditions incidentally learned responses are better retained than responses that have been intentionally repeated. The first question this raises is that of the nature of the connotative response, and the second is the reason for its seemingly unusual retention. As was indicated earlier, two alternative theories seem capable of providing answers to these questions. Our interpretation of Osgood (5) is that he regards the connotative meaning of a given stimulus as: (*a*) constituting some sort of behavioral pattern; (*b*) capable of giving rise to verbal associative judgments of its connotative quality; and (*c*) durable because of its inherent physiological nature. The present interpretation of Bousfield's theory (1) is that the connotative evaluation of a given stimulus obtained by means of the semantic differential is best explained by assuming it is: (*a*) essentially a simple verbal associative response; (*b*) elicited either directly by its given stimulus or mediationally through other associative responses that have become linked to the given stimulus; and (*c*) durable because of the factor of safety deriving from its mediation by multiple alternative associative responses. Both these theories seem to provide an explanation of the apparent resistance to forgetting of evaluative responses, and it would be desirable to test their relative merits. There are several problems in such an undertaking which seem to involve Osgood's theory more than that of Bousfield. Actually Osgood (5 p. 7) cautiously leaves open the question of the underlying nature of the mediating process representing meaning, and admits it may comprise purely neural events. Osgood's meaning, while described as behavioral in nature, is as yet unspecified. On the other hand, verbal responses, which Bousfield regards as meaningful, as specific and readily identified. Perhaps the best that can be done is to study the relative permanence of various types of conditioned responses. How permanent, for example, are conditioned motor, cardiac, and psychogalvanic responses? If these responses can be shown to outlast conditioned verbal responses, there might be some basis for preferring Osgood's theory. In any event this would be useful information. Another alternative study would deal with connotatively neutral verbal stimuli. The training phase here would involve the learning of paired associates composed of nonsense syllables coupled with the neutral meaningful words. The testing to be conducted after forgetting has taken place, would make use of the recognition method. A list of five or six words including one of the verbal associates of the response member of a pair would be presented to the subject whose task would be that of indicating the word which he would most readily associate with the nonsense syllable.

Research along these lines should yield results not only of interest from the standpoint of learning theory, but should also aid in the understanding of problems encountered in the study of personality and clinical psychology.

E. SUMMARY

This paper reports one of two related studies for testing the following experimental hypothesis: Subsequent to the learning of a given meaningful response presented as a translation of a paralog, there should be a tendency for the connotatively meaningful evaluative responses acquired by the paralog as a result of this learning to be retained when the verbal response itself can no longer be recalled.

By means of the method of paired associates, 25 subjects individually learned the purported single-word translations of 18 paralogs which were actually Turkish words. The translation words were selected from semantic differential norms so that six had Good, six had Bad, and six had evaluatively Neutral ratings. These words, furthermore, had approximately similar frequencies-of-use according to the Thorndike-Lorge tables. One week after the initial learning, the subjects were asked to recall the purported translations of the paralogs, and to rate the meanings of the paralogs on the good-bad semantic differential scale.

The analyses of the data indicated: (*a*) When errors were made in recall, the connotative values of the incorrectly recalled words tended to match the words they displaced. (*b*) When subjects produced no responses to the paralogs, they tended nevertheless to assign evaluative meanings to the paralogs which matched the values of the words with which the paralogs had been paired in the original learning. (*c*) The most poorly recalled translation words were those with Bad evaluative meanings, and it was found that these words had the smallest amount of meaningfulness as computed by Noble's method. The theoretical significance of these findings is considered from the standpoint of verbal learning theory.

REFERENCES

1. Bousfield, W. A. The significance of partial response identities for the problem of learning. Paper presented at the Annual Meeting of the American Psychological Association, Washington, D.C., Aug., 1958.
2. Jenkins, J. J., Russell, W. A., & Suci, G. J. An atlas of semantic profiles for 360 words. Tech. Rep. No. 15, ONR Contract N8onr-66216, 1957.
3. Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421-430.
4. ———. The role of stimulus meaning (m) in serial learning. *J. Exp. Psychol.*, 1952, **43**, 437-446.
5. Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. The Measurement of Meaning. Urbana, Ill.: Univ. Illinois Press, 1954.
6. Thorndike, E. L., & Lorge, I. The Teacher's Word Book of 30,000 Words. New York: Teachers College, Bureau of Publication, 1944.
7. Yavuz, H. S., & Bousfield, W. A. Recall of connotative meaning. *Psychol. Rep.*, 1959, **5**, 319-322.

Norwich Hospital
P. O. Box 508
Norwich, Connecticut

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 419-425.

LANGUAGE PROCESSING ABILITY OF PROCESS AND REACTIVE SCHIZOPHRENICS*[1]

Veterans Administration Hospital, Battle Creek, Michigan

DAVID PEARL

A. INTRODUCTION

Language can be considered as a vehicle for the expression and communication of thought. The development of language occurs in a matrix of social interaction. Through differential reinforcement, certain expectancies are developed with respect to word and sentence structure and the frequency of letter or word occurrences. Implicitly comprehended statistical regularities in letter and word combinations together with the information provided by context and the considerable redundancy present in language help prevent misunderstanding in communication and reduces error in the attribution of meaning. To the extent that psychopathological conditions impair subjects' thought processes through a loosening of associations and a breakdown of expectancies, the ability to utilize information from language redundancy and contextual meaning is reduced.

Language and thought disorders are considered as primary symptoms of schizophrenia. Cameron (3) has stressed the loss of communicability and that the essential schizophrenic disturbance lies in the social disarticulation of the patient. Studies by Baker (1) and Ellsworth (4) indicate that the language genetic level of adult schizophrenics is comparable with that of young children rather than normal adults. Manifestations of schizophrenics' language indicate concreteness of language attitudes (1), rigidity, repetitiousness (8) and approximation, fragmentation, and over-inclusiveness (7). Cameron points out that one finds a general defect in selection and elimination of competing terms in schizophrenic language (3).

The schizophrenic thus would differ from normals by deviations in the language response hierarchies which ordinarily are developed through a common language experience. The formal structure of schizophrenic language

---

* Received in the Editorial Office on December 21, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] A number of these subjects were participants in a Veterans Administration Chemotherapy Cooperative Study. The assistance of Thomas Eliseo of the VA Hospital, Lebanon, Penn. in collecting a portion of protocols, and the staff of the VA Central NP Research Laboratory, Perry Point, Md. in providing the abbreviated Elgin Scale is gratefully acknowledged.

would tend to be different and the schizophrenic should be less adept than normal subjects in utilizing information from language structure, redundancy, and context.

Schizophrenics vary widely in terms of prognosis and morbidity. Current classificatory schema stress life history as a criterion. The process-reactive dichotomy is widely utilized. Process schizophrenia is defined as a behavioral disturbance of gradual and insidious onset with a lack of clearly discernible stress factors often beginning in early childhood. Reactive schizophrenics on the other hand, are conceived as having a relatively normal prepsychotic personality adjustment and as experiencing a sudden onset of psychosis in the presence of precipitating factors. Prognosis is more favorable for reactive than process schizophrenics.

The long insidious development of illness throughout the formative years in the instance of process schizophrenics should interfere with the development of appropriate language techniques involved both in the sending and receiving of messages. With a deviant socialization experience accompanied by an autistic distortion of reality, the process schizophrenic should not have developed language response hierarchies generally common to members of the same linguistic community. This would not be the case for the reactive schizophrenic patient since his relatively good prepsychotic personality adjustment would have permitted the development of normal language skills and language stimuli hierarchies. The later onset of a schizophrenic reaction however could be expected to produce distortion in linguistic expectations for the reactive schizophrenic, particularly during the acute stages of the psychotic reaction. With partial stabilization accompanied by beginning restitution, the reactive schizophrenic should regain his prepsychotic linguistic framework. He should differ less from the normal adult in message sending and receiving than the process schizophrenic.

The hypotheses investigated in this study therefore were (*a*) schizophrenic patients would be inferior to normal subjects in message interpretation through the utilization of information offered through language redundancy, structure and context; and (*b*) reactive schizophrenics will demonstrate greater facility in message encoding than process schizophrenics.

B. Method

Shannon's guessing game technique (7) for determining the informational content of language sequences was adapted for use with a schizophrenic subject population. Shannon's method assumes that persons have an implicit knowledge of language statistics and when required to guess the next letter of a meaningful language sequence, would utilize this information to predict the most likely letter. If their first guess is incorrect, they continue guessing until correct. Then in a similar manner, predictions are made for successive letters of the language passage. As correct letters are established, uncertainty as to succeeding letters is diminished and the number of incorrect guesses is reduced. As utilized by Shannon, this technique required a subject to encode a series of passages in order to compile a representative frequency distribution of guesses for each letter position. In an analogous fashion for this study, numbers of guesses representative of a subject group were obtained through testing a number of subjects who were members of the group.

Two short sentences of approximately equal difficulty were selected for this study. These were: (*a*) the phone rang as he walked to the door, and (*b*) balance is important in becoming a dancer. These sentences were represented on separate paper sheets so that each letter of each word was represented by a dash and each word by a separate group of dashes. The structuring of the dashes into separate words was dictated by the need to avoid too great a level of difficulty for schizophrenic patients and because such groupings offered structural information whose utilization then could be determined.

Recently hospitalized schizophrenic patients were rated with an abbreviated Elgin Prognostic Scale (2). The rater's reliability had been previously established through correlating .84 with a second experience rater on another subject sample. Subjects with low and high Elgin scale scores were then classified as being reactive and process schizophrenics. Twenty-eight subjects thus rated as reactive schizophrenics and 40 patients classified as process schizophrenics were given the two sentences to encode. The same order was used throughout. For comparative purposes, 40 "normal" employees of the hospital were recruited and similarly tested.

Two measures were utilized. The first was the number of guesses required to encode both sentences while the second involved the number of guesses required to encode the last 17 letters of both sentences. It was assumed that the capacity to extract appropriate information from the already solved first half of the sentence and to utilize contextual information, would be measured through this second score.

It seemed obvious that the degree of utilization of language information would be a function of language facility. Therefore all subjects were given the Wechsler Adult Intelligence Scale (WAIS) Vocabulary subtest and scores therefrom served as measures of verbal proficiency.

C. RESULTS

Analysis of variance of process and reactive schizophrenics and normal subjects' vocabulary scores indicated that no significant difference in verbal facility was present between these groups. Nevertheless, subjects of each group were categorized as above or below the median vocabulary score for all subjects. Two-way analyses of variance were then calculated for both information utilization variables for the three subject groups and the level of verbal knowledge. Because of lack of proportionality of subject numbers for these classifications, random selection of 12 subjects for each diagnostic group and verbal level was made, thus providing data of 72 subjects for analysis. Hartley's F Max test demonstrated homogeneity of variance for both encoding variables.

Table 1 presents results of the analysis of variance for the encoding of both sentences and indicates that the several groups differ significantly in number of guesses required for sentence reconstruction. One-tailed *t* tests indicated that overall, normal subjects required significantly fewer guesses than did process or reactive subjects, at the *P*.01 and *P*.05 levels, respectively. Process and reactive groups however did not differ significantly. Subjects with a higher vocabulary level also required significantly fewer guesses than those with less verbal proficiency. A significant interaction between subject groups and vocabulary level was present. High vocabulary subjects from the three groups did not differ significantly and although these subjects did not differ from low vocabulary "normals," they required significantly fewer guesses than did low vocabulary schizophrenics. Both low vocabulary "normals" and reactive schizophrenics differed significantly from the low vocabulary process schizophrenics who were least adept. Significance levels are presented in Table 2.

TABLE 1
GUESSING SCORES FOR SENTENCES: MEANS AND ANALYSIS OF VARIANCE

	Means						
Subjects	High vocab.	Low vocab.	Source	*df*	MS	*F*	*p*
Process	67.1	151.7	Groups	2	8050	3.3375	$< .05$
Reactive	91.8	110.0	Vocabulary level	1	22969	9.5228	$< .01$
Normals	70.2	74.5	Interaction	2	12034	4.982	$< .01$
			Within groups	66	2412		

Results for the analysis of variance of the "last 17 letters" scores are presented in Table 3.

Groups differed significantly and *t* tests disclosed that "normals" required fewer guesses than either the process or reactive schizophrenics at the *P*.05

TABLE 2
SIGNIFICANCE LEVELS OF GROUP DIFFERENCES FOR SENTENCE ENCODING

	Process *S*s		Reactive *S*s		Normal *S*s	
	High vocab.	Low vocab.	High vocab.	Low vocab.	High vocab.	Low vocab.
High vocabulary process		*P*.001	NS	*P*.001	NS	NS
Low vocabulary process			*P*.01	*P*.001	*P*.001	*P*.001
High vocabulary reactive				NS	NS	NS
Low vocabulary reactive					*P*.05	NS
High vocabulary normal						NS

TABLE 3
GUESSING SCORES FOR LAST 17 LETTERS: MEANS AND ANALYSIS OF VARIANCE

	Means						
Subjects	High vocab.	Low vocab.	Source	*df*	MS	*F*	*p*
Process	39.5	54.8	Groups	2	1680	3.3600	$< .05$
Reactive	42.0	56.8	Vocabulary level	1	2301	4.5295	$< .05$
Normal	32.0	35.8	Interaction	2	252	.9902	NS
			Within groups	66	508		

and *P*.02 levels, respectively. No difference was found between the two schizophrenic groups. Again, higher vocabulary subjects required significantly fewer guesses at the *P*.05 level.

D. DISCUSSION

Results have therefore indicated that the level of verbal facility is generally a greater determinant of language information processing ability for these subjects than is the presence or absence of a psychotic process. However, only patients participated who were sufficiently in reality contact to be cooperative, thus possibly reducing the interference effects associated with more disturbed schizophrenic states.

When vocabulary levels were disregarded, process and reactive subjects did not differ significantly although both were inferior to "normal" subjects for both information processing variables employed. Hypothesis (*a*) relative to the differential encoding facility of "normals" and schizophrenic subjects is therefore supported. Hypothesis (*b*) however was not supported in a comparison of all reactive and process schizophrenics though a significant differential was present when low vocabulary process and reactive subjects were compared. Vocabulary level can be considered as an index of previous verbal learning ability. It may be inferred therefore that process schizophrenics with greater vocabularies experienced less interference in the development of language skills and linguistic expectations than did low vocabulary level process subjects. Possibly the psychotic process did not evolve for greater vocabulary process schizophrenics until after language stimuli hierarchies and linguistic expectations were developed. Such expectations however would have been impaired or erratically developed for low verbal facility process schizophrenics since these subjects were significantly poorer in sentence encoding than were reactive schizophrenics with similar vocabulary command.

Since no difference was found between these low vocabulary psychotic groups in encoding the last 17 letters of the sentences, it appears that these groups differed with respect to the initial approaches to sentence encoding. A qualitative analysis of guesses for the first halves of sentences indicated that though erratic hypotheses regarding letter expectancies characterized both groups, the low vocabulary process schizophrenics persevered longer with random trial and error guesses. Completely inappropriate guesses were often offered by these subjects indicating difficulty in evaluating letter probabilities and a failure to take advantage of word context. Low vocabulary reactive schizophrenics however offered fewer random guesses, were somewhat more systematic in eliminating inappropriate responses and were more clearly cognizant of word and phrase structure.

The effects of chemotherapy on the functioning of schizophrenics is of considerable interest. To the extent that it might foster increased reality orientation, chemotherapy might be expected to increase schizophrenics' ability to utilize linguistic information. In an associated study (6), 29 of the schizophrenic subjects participating in the investigation just described, were given two additional sentences to encode following a three-month period of chemotherapy. These additional sentences were also administered to this study's "normal" control subjects. The median number of guesses required by the "normal" group for each letter position served as a standard. The performance of each schizophrenic subject for the two pre- and two post-chemotherapy sentences was expressed as a deviation from the "normal" median. The algebraic summation of a subject's deviations served then as his deviation score. Improvement in language processing capability was defined as a shift toward the "normal" group median as indicated by smaller post-chemotherapy deviation scores. Findings indicated that with vocabulary level controlled, phenothiazine or reserpine tranquilizers failed to induce significant improvement.

E. SUMMARY

The Shannon guessing game technique was utilized to study the ability of process and reactive schizophrenic subjects to utilize information available in the redundancy and structure of the English language. The number of guesses required to encode two sentences was compared for reactive and process schizophrenic groups with a "normal" control group. Verbal facility was controlled through a dichotomization of subjects through WAIS Vocabulary subtest scores into those of lesser and greater vocabulary skill. Analyses indicated that for both schizophrenic groups, the level of vocabulary facility is a greater determinant of language information processing ability than the presence of psychosis. Low vocabulary process schizophrenics however were significantly inferior in performance than low vocabulary reactive schizophrenics and both of these groups were inferior to "normal" subjects with a similarly low vocabulary skill.

REFERENCES

1. BAKER, R. W. The acquisition of verbal concepts in schizophrenia: A developmental approach to the study of disturbed language behavior. *Dissertation Abstr.*, 1953, **13**, 870.
2. BECKER, W. C. The process-reactive distinction: A key to the problem of schizophrenia? *J. Nerv. Ment. Dis.*, 1959, **129**, 442-449.
3. CAMERON, N. Reasoning, regression and communication in schizophrenics. *Psychol. Monog.*, 1938, **50**, No. 1.
4. ELLSWORTH, R. B. The regression of schizophrenic language. *J. Consult. Psychol.*, 1951, **15**, 387-391.
5. MABRY, M. A study comparing the language characteristics of scattered and non-scattered schizophrenics with normals. *Dissertation Abst.*, 1955, **15**, 457-458.
6. PEARL, D. Unpublished study, 1961.
7. SHANNON, C. E. Prediction and entropy of printed English. *Bell Syst. Tech. J.*, 1951, **30**, 50-64.
8. WHITE, M. A. A study of schizophrenic language. *J. Abn. & Soc. Psychol.*, 1949, **44**, 61-74.

Mental Health Study Section
Division of Research Grants
National Institutes of Health
Bethesda 14, Maryland

[illegible] scores. Findings indicate the [illegible]

[illegible]

SUMMARY

[illegible]

BIBLIOGRAPHY

[illegible]

[illegible]
[illegible]
National Institutes of Health
Bethesda 14, Maryland

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 427-436.

THE EFFECT OF VOLUNTARILY DIRECTED ATTENTION ON SUCCESSIVE CORTICAL ACTIVATION RESPONSES*

Veterans Administration, Bedford, Massachusetts

THOMAS MULHOLLAND AND SYLVIA RUNNALS

A. INTRODUCTION

Though it is well known that electroencephalographic (EEG) changes are associated with changes in consciousness (1), the study of the temporal course of the attention process, as it is directed voluntarily, remains a challenging methodological as well as theoretical problem. Usually psychological events such as voluntary attention cannot be matched to simultaneous electroencephalographic responses with much assurance because evaluation of the time of occurrence of psychological events is often uncertain.

In this study, a method was applied which permitted a more definite association between the occurrence of a psychological process and the occurrence of an electroencephalographic cortical activation response. This method employed a stimulus-brain response feedback loop in conjunction with special instructions to the subject (2, 4), and confirms and extends a previous study of voluntarily directed attention using the "feedback" method (2).

The feedback loop was based on the suppression of the EEG alpha rhythm in response to a stimulus. Such a loop is created when the stimulus which causes alerting is automatically produced whenever alpha (at or above a critical amplitude) occurs; when alpha is suppressed or attenunated enough, the stimulus is automatically removed, etc. Under these conditions a series of cortical activation responses occurs associated with a series of stimulus events. The subject can perform various cognitive tasks such as counting the successive stimuli, "paying attention" to a particular stimulus event, etc. In this way the effect of complex cognitive tasks may be evaluated by studying their effect on the series of cortical activation responses.

B. EXPERIMENT 1

1. *Method*

a. Apparatus. The feedback loop, like that described previously was composed of (*a*) a Grass Model V Polygraph to detect and amplify the

---

* Received in the Editorial Office on December 21, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Presss.

brain rhythms; (*b*) a variable frequency, variable selectivity filter of the brain rhythms; (*c*) an amplifier of filter output; (*d*) a relay controlled by (*c*) and controlling (*e*) a lamp; (*f*) a human subject. The stimulus was produced by an incandescent lamp located 10 feet from the subject. When the lamp was *ON* it did not flicker. Incident illumination at the eyes was about eight foot candles. The lamp could be switched *OFF* independently of the stimulus relay. This permitted evaluation of system response in the absence of the stimulus (2, 3, 4, 5).

b. Procedure. Twenty-two volunteers (men and women) were each tested under each of four attention conditions: *A,* simple viewing without counting; *B,* counting to the 15*th* stimulus event followed by no counting; *C,* counting to the 30*th* stimulus event; *D,* counting to the 15*th* event and judging whether it was about the same or different brightness as the events before it, followed by no counting. All counting and judging tasks were performed silently, the subject being instructed to "mentally count"; to make a "mental" judgment, etc.

Consistent with previous studies, testing was done in a sound-reduced, air conditioned, dark room. The subject was seated comfortably with eyes closed. (The stimulus was always presented through closed lids to prevent unnecessary tearing or irritation). Recordings were obtained using bipolar scalp electrodes placed over parietal-occipital regions on one side.

The duration of the testing session was about one half hour or less. A brief resting record was obtained to determine the dominant alpha frequency. The filter was tuned to that value and the threshold of stimulus occurrence was adjusted until the stimulus relay went *ON* when alpha was present and *OFF* when alpha was not occurring. For some subjects having prominent alpha rhythms the usual "blockade" of alpha did not occur with increased attention. For these, alpha was attenuated during attention. This required the stimulus threshold be so adjusted that the stimulus would go *ON* at the range say, between 50–100 per cent of resting alpha amplitude and go *OFF* below that value. Thus the stimulus threshold was not necessarily the same for each subject. After a period of quiet resting, yielding evident and persistent alpha, the lamp was switched *ON* and feedback stimulation continued until about thirty successive activation responses had occurred.

2. *Results*

The average durations of relay *ON* and *OFF* events were obtained for 10 events before and 30 events during feedback stimulation. The stimulus durations during feedback were analyzed using analysis of variance.

The event durations for the period of feedback were not distributed as a Gaussian function and were normalized using the Freeman-Tukey transformation and analysis of variance applied (*ON* and *OFF* events were analyzed separately). Variance was partitioned among Sequence, Serial Position, Attention conditions and the interaction of these. A summary of the analyses is presented in Table 1. As can be seen, the *F* tests associated with Sequence, Serial Position, Sequence × Position, and Sequence × Position × Attention were significant. A similar analysis is shown for *OFF* durations. The *F* tests associated with Serial Position, Attention, and Sequence by Attention were significant.

TABLE 1
SUMMARY ANALYSIS OF VARIANCE (*ON* AND *OFF* DURATIONS)

Source of variance	*df*	*ON* Mean square	*F*	*p*	*OFF* Mean square	*F*	*p*
Sequence (Sq)	3	2763.3	3.3^1	$<.05$	1103.0	1.5^1	
Subjects within Sq	18	846.9			740.3		
Serial position (P)	29	48.1	4.9^2	$<.001$	24.1	2.4^2	$<.001$
Attention (A)	3	61.0	0.7^3		124.4	7.8^3	$<.001$
P × A	87	6.5	1.1^4		11.3	1.1^4	
Sq × P	87	15.8	1.6^2	$<.01$	9.6	$<1.0^2$	
Sq × A	9	54.2	0.6^3		52.0	3.3^3	$<.01$
Sq × P × A	261	9.0	1.5^4	$<.01$	9.6	0.9^4	
Pooled S × P	522	9.9			10.0		
Pooled S × A	54	85.4			15.9		
Residual	1566	6.0			10.4		
Total	2639						

[1] Tested against subjects within Sq.
[2] Tested against Pooled S × P.
[3] Tested against Pooled S × A.
[4] Tested against Residual.

Since the Serial Position × Attention interaction was not significant another analysis was made of the 10*th* through the 20*th* feedback event durations. The summary is presented in Table 2. For *ON* durations *F* tests associated with Sequence and Sequence × Position × Attention were significant; for *OFF* the *F* tests associated with Attention, Serial Position × Attention, and Sequence × Attention were significant. The significant interactions involving the sequence variable indicate that the effects of attention and the change over series of activation events are a function of the particular location of an experimental condition in a sequence of conditions. The Serial Position × Attention interaction indicates the variation of serial events as a function of the particular attention condition. The average durations of stimulus relay *ON* and *OFF* obtained before and during

TABLE 2
SUMMARY ANALYSIS OF VARIANCE (*ON* AND *OFF* DURATIONS 10–20)

Source of variance	*df*	*ON* Mean square	*F*	*p*	*OFF* Mean square	*F*	*p*
Sequence (Sq)	3	1390.0	3.1[1]	=.05	407.1	1.4[1]	
Subjects within Sq	18	449.7			292.8		
Serial position (P)	10	6.7	<1.0[2]		17.7	1.4[2]	
Attention (A)	3	28.5	0.6[3]		118.8	10.1[3]	<.001
P × A	30	6.9	1.1[4]		14.5	1.5[4]	=.05
Sq × P	30	5.7	0.8[2]		13.5	1.1[2]	
Sq × A	9	45.6	0.9[3]		25.3	2.1[3]	=.05
Sp × P × A	90	8.2	1.3[4]	<.05	11.7	1.2[4]	
Pooled S × P	180	6.8			12.8		
Pooled S × A	54	51.1			11.8		
Residual	540	6.3			10.0		
Total	967						

[1] Tested against subjects within Sq.
[2] Tested against Pooled S × P.
[3] Tested against Pooled S × A.
[4] Tested against Residual.

feedback stimulation for each attention condition are presented in Figure 1.

The changes over successive events before and during feedback occurred for both system *ON* and *OFF* events. For expository convenience the difference between the series of *ON* and *OFF* events is used here to describe the results as shown in Figure 1. Statistical validation is provided in Tables 1 and 2 for each event series separately, not for the difference between the two event series.

During the resting condition *ON* durations (associated with the duration of alpha bursts) are longer than the durations of *OFF* events (associated with periods wherein alpha was not occurring or was occurring below a critical amplitude). This indicates that as the difference (*OFF* minus *ON* durations) becomes less positive or more negative, system behavior resembles that obtained in the resting condition.

With the onset of feedback stimulation, the difference between *OFF* and *ON* durations is reversed from that obtained in the resting condition, i.e., the system *OFF* responses (associated with cortical activation responses) become longer and *ON* responses (associated with alpha bursts) become shorter. During Condition *A* (simple viewing) this reversal is definite for the first eight events since *OFF* durations decrease and *ON* durations increase as a function of successive events.

For Condition *B* (counting to 15*th* event) reversal of the resting pattern

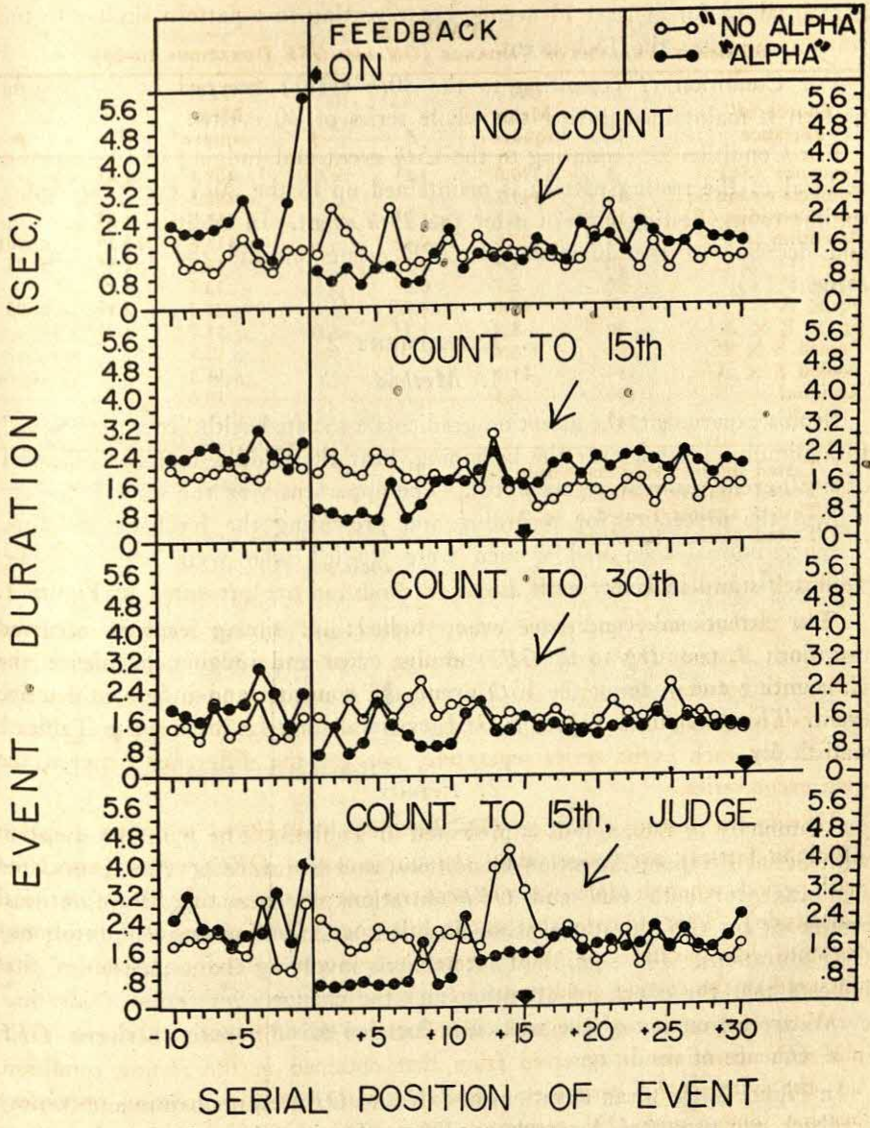


FIGURE 1

Series of *ON* ("alpha") and *OFF* ("no-alpha") mean durations for twenty-two normals during four attention conditions, before and during feedback stimulation. Experiment 1.

is maintained for at least 11 events, but reverting to a pattern similar to the resting one after the 15th event.

For Condition *C* (counting to the *30th* event) reversal of the resting pattern is maintained during the whole series of 30 events.

For Condition *D* (counting to the *15th* event and judging its brightness), reversal of the resting pattern is maintained up to the *20th* event, becoming similar to the resting pattern after the *25th* event. In addition, an increase and decrease of the durations of *OFF* events around the *15th* event is evident (2).

C. EXPERIMENT 2

1. *Method*

In this experiment, the attention gradients associated with "counting-down" to a stimulus located near the beginning, near the middle, and near the end of a 30-event series were evaluated. The apparatus was the same as before as was the procedure for recording and presenting the feedback stimulus. Twenty normal men and women were studied who were not the same subjects studied in Experiment 1.

Four attentional conditions were studied: *A,* simple viewing without counting; *B,* counting to the *7th* stimulus event and judging its brightness; *C,* counting and judging the *13th* event; *D,* counting and judging the *23rd* event. Thirty stimulus *ON* and *OFF* events during feedback were analyzed statistically.

2. *Results*

A summary of the analysis is presented in Table 3. The *F* ratio associated with Serial Position, Attention Conditions, and Sequence × Attention were significant for both *ON* and *OFF* durations; Sequence × Position was significant for *ON* durations, while Position × Attention was significant for *OFF* durations. The significant interactions involving the sequence variable indicate that the effects of attention and the change over series of alerting events are a function of the particular location of an experimental condition in a sequence of conditions.

In Figure 2, the mean durations of *ON* and *OFF* events before and during feedback are presented for each condition. During *A,* no definite gradient was associated with any particular event; for Condition *B,* the *OFF* durations showed a definite increase at the *6th OFF* event followed by a rapid decrease of *OFF* durations; during Condition *C,* an increase of *OFF* durations occurred around the *11th OFF* event, followed by a decrease of *OFF* du-

TABLE 3
SUMMARY ANALYSIS OF VARIANCE *ON* AND *OFF* DURATIONS

Source of variance	*df*	*ON* Mean square	*F*	*p*	*OFF* Mean square	*F*	*p*
Sequence (Sq)	3	401.9	1.2^1		756.7	2.2^1	
Subjects within Sq	16	336.7	—		347.2	—	
Serial position (P)	29	10.6	2.7^2	$<.001$	40.1	4.3^2	$<.001$
Attention (A)	3	59.8	6.0^3	$<.01$	108.0	6.9^3	$<.001$
P × A	87	4.0	1.1^4		13.3	1.4^4	$<.05$
Sq × P	87	6.4	1.6^2	$<.01$	9.4	1.0^2	
Sq × A	9	42.5	4.3^3	$<.001$	40.2	2.6^3	$<.05$
Sq × P × A	261	3.4	$<1.0^4$		9.5	$<1.0^4$	
Pooled S × P	464	3.9	—		9.4	—	
Pooled S × A	48	9.9	—		15.6	—	
Residual	1392	3.8	—		9.6	—	

[1] Tested against subjects within Sq.
[2] Tested against Pooled S × P.
[3] Tested against Pooled S × A.
[4] Tested against Residual.

rations. This gradient did not seem to be as sharply peaked as in Condition *B;* in Condition *D* no definite gradient can be noted. However, the durations of *OFF* events were longer here than in the condition of simple viewing, or during the last 15 *OFF* events in Conditions *B* and *C*.

ON durations showed no definite gradients associated with any of the stimulus events. They were generally less during the count-down conditions than during simple viewing *A*. This effect, however, was much more evident for Condition *D* than for the other conditions.

D. DISCUSSION

These results confirm the previous experiment showing the effect of voluntarily directed attention on successive cortical activation responses as they are produced in the feedback loop described here (2). The three experiments taken together show that internal, private, subjective, voluntarily directed attention states, which are induced by prior instruction and activated by the subject, can be reliably inferred from the changes of feedback system behavior.

Since the *ON* and *OFF* behavior of the system is partially determined by the occurrence of alpha bursts and alpha "blockade" or cortical activation, it can be hypothesized that subjective, voluntarily directed attention states (induced by a communication, remembered by the subject and activated by the subject) can so modify the durations of cortical activation responses that a useful monitoring of subjective, voluntarily directed attention processes

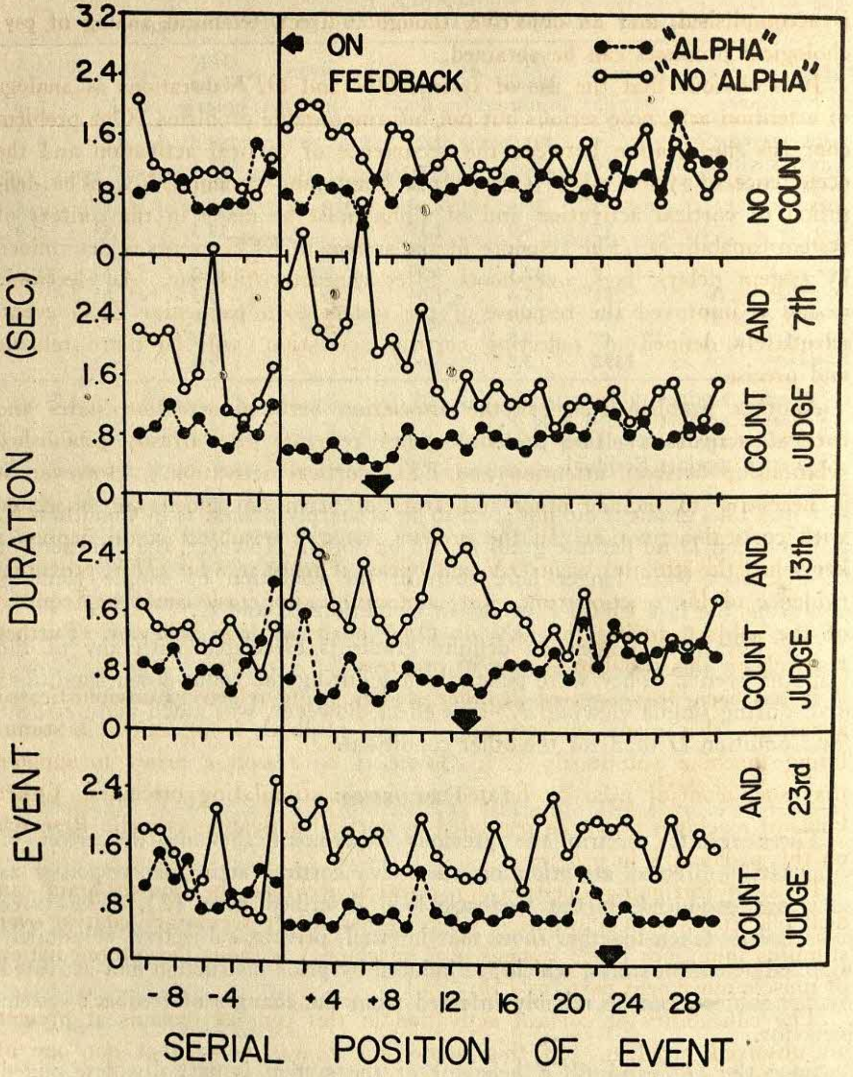


FIGURE 2
Series of *ON* ("alpha") and *OFF* ("no-alpha") mean durations for twenty normals during four attention conditions, before and during feedback stimulation. Experiment 2.

is accomplished, and an objective, though indirect, temporal analog of psychological processes can be obtained.

It is obvious that the use of system *ON* and *OFF* durations as analogs of attention acts, pose serious but not insurmountable problems. One problem concerns the relation between the occurrence of cortical activation and the occurrence of system *OFF* (also, alpha bursts and system *ON*). The definition of cortical activation and of alpha must be made in the context of system capabilities. The response of the system to EEG events is determined by system delays, lags, overshoots, filter characteristics, etc. As electronic design is improved the response of the system to a particular EEG event, adequately defined as reflecting cortical activation, will be more reliable and precise.

Another problem concerns the association between attention states and cortical activation. The present studies reiterate the already established relationship between attention and EEG cortical activation. However, it is necessary to include other indicators of attention which are associated with cortical activation. In the present system the subject could depress a key when the stimulus went *ON* and release it when it went *OFF*, providing evidence of his reaction time. Or, a discrimination task could be required of the subject during the *ON* or *OFF* phase of loop behavior. Further research on this problem is now in progress.

It has been demonstrated in other studies, using relatively unsophisticated apparatus, that many subjects can, via the feedback loop, control a stimulating machine voluntarily (5). There is no reason *a priori* to suppose that such control need be limited to simple stimulating machines. Other kinds of machines might be regulated by cortical activation as well. Research on this problem is in progress.

It is an intriguing question for research whether the human brain can deal efficiently with a situation where it has direct regulatory control over stimulus input (or other systems, e.g., machines) without the intermediation of muscle movement patterns (5).

The educability of cortical activation in this context remains at present an unsolved problem. (In this connection it was noted that not one of over 150 subjects whose brain reactions were part of the feedback system ever reported that they realized that something *they* did was controlling the occurrence of the light).

E. SUMMARY AND CONCLUSIONS

With a "feedback" loop, frequencies in the alpha range caused a stimulus automatically to occur. When alpha was suppressed the stimulus was automatically removed. When alpha reoccurred the stimulus was automatically presented again, etc. During feedback stimulation the following phenomena were observed: (*a*) the temporal pattern of system *ON* and *OFF* durations was significantly different for conditions of viewing and silently counting feedback stimuli compared to simple viewing when no such counting was required. In general, *ON* durations were shorter, *OFF* durations longer and the variance of *ON* durations reduced during the conditions requiring counting. (*b*) Internally directed attention gradients (produced by voluntarily directing attention to specific stimuli) were detected and evaluated. (*c*) It was noted that this system permits regulation of a stimulating machine by cognitive acts, and that extension of this control to other external systems would depend on the educability of volitional control of brain alerting.

REFERENCES

1. LINDSLEY, D. B. Attention, consciousness, sleep and wakefulness. *In* Field, J., Magoun, H. W., & Hall, V. E. (*Eds.*), *Handbook of Physiology,* Section 1: Neurophysiology, Vol. III, American Physiological Society, Washington, D.C., 1960, 1553-1593.
2. MULHOLLAND, T. The electroencephalogram as an experimental tool in the study of internal attention gradients. *Transactions of the N.Y. Academy of Sciences,* 1962, **24**, 664-669.
3. MULHOLLAND, T., & RUNNALS, S. A brain response-sensory stimulus feedback system. *In:* Digest of Papers, 1961 International Conference on Medical Electronics, New York City. Publisher: Conference Committee for 1961 International Conference on Medical Electronics, L. E. Flory Chm., Princeton, N.J., 1961, p. 166.
4. ———. A stimulus-brain feedback system for evaluation of alertness. *J. of Psychol.,* 1962a, **54**, 69-83.
5. ———. The evaluation of alerting with a stimulus-brain feedback loop. *Electroenceph. Clin. Neurophysiol.,* 1962b, **14**, 847-852.

VA Hospital
200 Springs Road
Bedford, Massachusetts

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 437-440.

A STUDY OF SELF-DISCLOSURE IN THE CLASSROOM*

Department of Psychology, University of Arkansas

PHILIP HIMELSTEIN AND WILSON W. KIMBROUGH, JR.[1]

A. INTRODUCTION

Psychologists and other behavioral scientists have long been interested in the procedures by which individuals make themselves known to others. Clinical and industrial psychologists have been concerned with those aspects of the self as revealed in interviews, whether for therapeutic or personnel purposes. As Jourard (2) points out, little is known about the self-disclosing process as it unfolds in everyday, unstructured situations. Yet, it is obvious that individuals do reveal much of themselves to others in social and other informal contacts.

One procedure for disclosing material about the self is the introduction, in which the individual is asked to say something about himself to a group of others who are, for the most part, strangers. This procedure is employed by many college instructors on the first meeting of a class as a method for developing an informal classroom atmosphere and reducing tension. In a typical situation employing the self-introduction procedure, each student gives his name and any other information about the self that he would care to reveal. Here, then, is an ideal setting for exploring the process of self-disclosure in a group setting.

The purpose of this paper is to investigate this process of self-disclosure in a classroom situation. Specifically, this study is designed to explore the possibility that a questionnaire designed to measure self-disclosure in the past is predictive of the amount of information that in individual reveals in a self-introduction situation, and of the time spent in introducing one's self. The study was also designed to explore the possibility that knowledge of social factors would be more predictive of self-disclosing behavior than would psychometric procedures.

---

* Received in the Editorial Office on December 28, 1962, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors would like to express their appreciation to Robert H. Hickson for a critical reading of the manuscript and to Glen A. Cole, director of the University of Arkansas Counseling and Guidance Training Institute, for his cooperation in the conduct of this study.

B. METHOD

1. *Subjects*

Twenty-five graduate students in education were employed as subjects for this study. All of the students were secondary school counselors participating in a counseling and guidance training institute supported by the National Defense Education Act, and conducted on the campus of the University of Arkansas during the summer of 1962.

2. *The Questionnaire*

Jourard has devised a questionnaire designed to measure the extent to which individuals disclose themselves to different target persons (mother, father, same-sex friend, opposite-sex friend or spouse). In addition to scores for specific target-persons, the scale provides an overall disclosure score. Jourard has employed this instrument in a wide variety of studies (1, 3) and, for detailed information concerning scoring the reader is referred to these publications. Previous research with this instrument (1) revealed an odd-even reliability coefficient of .93 for the total score.

3. *Procedure*

The entire procedure was conducted on the first day of class. During the first hour, the director of the institute asked each individual to stand and introduce himself. Beyond asking each student to give his name and to include his present professional position, no other instructions were given. Each individual was called upon for self-introductions in alphabetical order.

While the students in the class were going through the process of self-introductions, two raters, seated in the rear of the classroom, scored each student for the amount of self-information revealed. This was performed by a check-off procedure in which check-marks were made for each item of information revealed (marital status, previous employment, military service, etc.). An individual's score for each rater was the total number of check-marks received. At the same time, each student was timed by each rater for total time spent on the introductions. Later the same day, the self-disclosure questionnaire was administered to the entire group with the comment that the instrument was for "research purposes only."

C. RESULTS AND DISCUSSION

Inter-rater reliability for the amount of information revealed in the self-introduction situation, calculated by the rank-order method, was .75. While the correlation is significant beyond the .01 level, it does indicate that the score for amount of information revealed in this situation is a fallible criterion and that correction for attenuation would be appropriate. Rater agreement on the time score proved to be sufficiently reliable (rho = .97) to forego the necessity for correction for attenuation.

The correlation between total score on the self-disclosure questionnaire and the amount of information revealed was .102. Correction for attenuation did not raise the correlation coefficient to a statistically significant level. The correlation between questionnaire score and the time score was —.01. It is apparent that one cannot predict from the questionnaire, which purports to measure how much the individual has revealed in the past to specific target-persons, the amount of information that the individual will reveal in the situation described in the present study.

There is evidence in the present study that an understanding of the situation and social variables enables one to make more accurate predictions concerning an individual's performance in a self-introduction situation than knowing how much the individual has revealed about himself in the past. The results in the present study indicate that individuals who are called later in the introductory session reveal more about themselves, and take more time for the introductions. Rank-order correlations were computed between subjects' order of appearance and the amount of information revealed in the classroom, and time spent on this process. For this purpose, the students were rank-ordered from the most to the least amount of information revealed in the classroom and for the time score.

In this analysis, the rank-order correlation between order of appearance and amount disclosed was —.425 after correction for attenuation (the uncorrected rho was —.37). This is significant at the .05 level and indicates that there is a significant trend for the individual to reveal more information about the self as more individuals precede him in the self-introduction situation. The rank-order correlation between order of appearance and the time score was —.503, significant at the .01 level. Again, there is a significant trend for the individual to spend more time in self-disclosure as more individuals precede him in this situation.

The results of this study would seem to suggest that predictions concerning self-disclosing behavior can be made more accurately by control of the situational factors than by the use of psychometric devises. While knowledge of the individual's order of appearance is helpful in this respect, it does leave much of the variance in self-disclosing behavior unexplained. The results of this study, however, would lead one to suspect that this variance can be better accounted for by stricter control of the situational variables than by the development of more refined questionnaire procedures.

D. SUMMARY

Twenty-five graduate students in education were called upon to introduce themselves during the first class meeting. Subjects were rated for amount of information revealed in the introductions and for time spent on introductions. A questionnaire designed to measure self-disclosure was later administered to the class. Neither scores for amount of information revealed nor time scores were found to be significantly related to scores on the questionnaire. It was found, however, that both classroom-disclosure and time score were related to the order of appearance of the student.

REFERENCES

1. JOURARD, S. M. Self-disclosure scores and grades in nursing college. *J. Appl. Psychol.,* 1961, **45**, 244-247.
2. ————. Age trends in self-disclosure. *Merrill-Palmer Quart.,* 1961, **7**, 191-197.
3. JOURARD, S. M., & LASAKOW, P. Some factors in self-disclosure. *J. Abn. & Soc. Psychol.,* 1958, **56**, 91-98.

Department of Psychology
University of Arkansas
Fayetteville, Arkansas

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 441-444.

FAKABILITY OF THE GORDON PERSONAL INVENTORY: REPLICATION AND EXTENSION*

Department of Psychology, University of Bridgeport

John R. Braun

A. Introduction

The Gordon Personal Inventory (GPI) is a forced-choice instrument containing scales for Cautiousness, Original Thinking, Personal Relations, and Vigor (3). A Total score which indicates the extent to which the individual has selected complimentary rather than derogatory alternatives may also be obtained. Little validation evidence is available although Braun, Alexander, and Weiss (1) found significant relationships between GPI scores and peer ratings of the corresponding traits.

Fakability was investigated by Braun (2) who had 15 male and four female industrial psychology class students complete the GPI under normal instructions and immediately thereafter under instructions to fake so as to appear the best possible candidate for an executive position in top management. With the exception of Personal Relations, scores on all variables were significantly increased under the faking set. The study suggested the possibility of using total score as a validity scale to detect faking, with either the 88th or 92nd percentiles as cutoff points.

The present investigation replicated and extended this work. This time *S*s faked "bad" as well as "good," and some *S*s were instructed to try to fake in such a way that their faking could not be detected.

B. Method

*S*s were male and female university students in various psychology courses. Three groups were used: Group I, undergraduate educational psychology class, fake-bad set, appear to be the poorest possible candidates for the top management executive position, $N = 28$ (23 female, five male); Group II, undergraduate educational psychology class, fake-good set, as in original investigation, $N = 24$ (21 female, three male); Group III, graduate course in psychological and educational testing—a service course for persons in education, fake-good set, with additional stipulation to try to fake in

---

* Received in the Editorial Office on January 13, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

such a way that we could not detect their faking, $N = 25$ (16 female, nine male).

All groups had normal administration first and faked administration second. Wilcoxon paired-replicates tests were used to assess the significance of the differences between scores from the two administrations.

C. Results and Discussion

Means and standard deviations are presented in Table 1. Groups II and III which had the fake-good sets showed significant increases for all scales except the Personal Relations one. This is completely consistent with the earlier study. Group I which had the fake-bad set showed significant decreases in score for all scales.

As in the earlier study, *S*s in the two fake-good groups made high scores on Cautiousness. The Group II faked mean of 9.58 would have a percentile

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR REGULAR AND FAKED ADMINISTRATIONS

	Group		
Variable	I	II	III
Cautiousness			
M (regular)	1.29	4.75	3.76
SD (regular)	8.02	6.66	6.74
M (faked)	— 8.25*	9.58*	8.20**
SD (faked)	3.43	3.89	2.71
Original Thinking			
M (regular)	3.96	3.92	5.24
SD (regular)	3.92	5.82	6.02
M (faked)	— 9.64*	9.12*	11.72*
SD (faked)	3.98	3.10	3.45
Personal Relations			
M (regular)	2.11	4.08	5.04
SD (regular)	6.68	4.91	7.03
M (faked)	— 5.57*	4.00	7.32
SD (faked)	3.74	3.51	2.68
Vigor			
M (regular)	2.46	3.67	3.36
SD (regular)	5.74	4.32	6.80
M (faked)	—10.46*	12.38*	10.20*
SD (faked)	4.51	3.44	3.05
Total			
M (regular)	9.82	16.42	17.40
SD (regular)	14.31	14.25	18.65
M (faked)	—33.92*	35.08*	37.44*
SD (faked)	9.13	6.02	2.23

* Significantly different from regular mean at .01 level.
** Significantly different from regular mean at level between .02 and .01.

equivalent between 82 and 86, while that for the Group III faked mean of 8.20 would be between 77 and 82. It is again suggested that in making these high scores on Cautiousness *S*s may have faked to appear more like middle than top management people (2, 4).

For the two groups with fake-good sets, mean faked total scores were 35.08 and 37.44. These mean scores place both groups beyond the 92nd percentile in terms of tendency to choose favorable rather than unfavorable alternatives on the inventory. This is true whether male or female college norms are used.

Attempting to use total score as a validity scale, if we said that any total score at the 92nd percentile or higher indicated faking, and any lower score indicated non-faking, we would be correct in 40 of 48 predictions for Group II. This would include six false negatives and two false positives. Using the 88th percentile as cutoff we would be correct in 43 of 48 predictions, with three false negatives and two false positives.

For Group III these figures would be 41 of 50 predictions with one false negative and eight false positives, and 41 of 50 with zero false negatives and nine false positives for the 92nd and 88th percentiles respectively. It is of interest that faking by Group III could be detected. *S*s in this group were instructed to answer so as to appear the best possible candidates, but also to answer so that the faking could not be detected. They were not able to satisfy both of these stipulations.

Taken together, the two investigations demonstrate clearly that *S*s can markedly alter their scores on the GPI when motivated to do so. They are able to increase or decrease their scores, depending on the particular set given them. The use of total score as a validity key to detect faking still seems promising however. Cutoffs established in the earlier investigation held up well with the new samples.

D. SUMMARY

An earlier investigation of the effects of a top management faking set on the Gordon Personal Inventory was replicated and extended. Three groups took the GPI under normal instructions and under various faking instructions. Involved were sets to fake-good, fake-good in such a way that faking could not be detected, and fake-bad. Earlier results were confirmed in that scores on most scales were significantly altered under the faking instructions. Also supported was an earlier suggestion that total score might be useful as a validity key to detect faking.

References

1. **Braun, J. R., Alexander, S. & Weiss, R.** Relationship between the Gordon Personal Inventory and consensus of peer ratings. *Psychol. Rep.*, 1961, **9**, 455.
2. **Braun, J. R.** Effects of a top management faking set on the Gordon Personal Inventory. *Psychol. Rep.*, 1962, **10**, 611-614.
3. **Gordon, L. V.** Manual, Gordon Personal Inventory. Yonkers, N.Y.: World Book, 1956.
4. **Porter, L. W., & Ghiselli, E.** The self perceptions of top and middle management personnel. *Person. Psychol.*, 1957, **10**, 397-406.

Department of Psychology
University of Bridgeport
Bridgeport, Connecticut

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 445-456.

PERSONALITY FACTOR STABILITY FOR THREE ABILITY LEVELS*[1]

System Development Corporation, Santa Monica, California

GERALD H. SHURE AND MILES S. ROGERS

A. INTRODUCTION

Population standardization data are considered essential for establishing the range of applicability of generalizations from psychometric tests. Strangely enough, this requirement is paralleled by rather little attention to the effect of systematic restriction in population characteristics on the factorial composition of tests and on the stability of factors and their interpretation. Perhaps the realization that there are many sources of variation in factor patterns other than that associated with population differences is responsible for this neglect. Yet, despite the interpretive problems posed by cross-sample comparisons of factors, a number of such studies have yielded fruitful results. Factor analysis of cognitive and personality tests for populations differing in age (1, 7) and neuropsychiatric classification (2, 12) indicates that for different populations some scales vary sufficiently in their factor composition to require important shifts in interpretation. In addition, the interpretable patterns of factor stability and instability found for the different populations help to clarify test interpretation.

Though many personality tests such as the California Psychological Inventory (CPI) are used with school, college, and industrial populations where wide ranges of intellectual ability are sampled, the effect of ability differences on the factor composition of these test scales has not been explored. Considering the extensive use of the CPI in which invariance is implicitly assumed, the question deserves an empirical answer. The patterns of factorial variance and invariance associated with intellectual ability not only have implications for counseling practices, but may provide important clues with respect to those personality dimensions which are stable or which should be redefined as a function of intellectual level.

The CPI is composed of 18 separate scales which have been grouped into

---

* Received in the Editorial Office on January 14, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] The authors wish to express their appreciation to Arnold Binder, Carles Cuadra, Harry Harman, and Burton Wolin who commented on an earlier draft of this paper and to Mary Taylor who exercised continuous editorial surveillance.

four areas of "adjustment and development" by Gough (5): I. Measures of poise, ascendancy and self-assurance include scales for Dominance (*Do*), Capacity for Status (*Cs*), Sociability (*Sy*), Social Presence (*Sp*), Self-Acceptance (*Sa*), and Sense of Well-Being (*Wb*). II. Measures of socialization, maturity and responsibility include scales for Responsibility (*Re*), Socialization (*So*), Self-Control (*Sc*), Tolerance (*To*), Good Impression (*Gi*), and Communality (*Cm*). III. Measures of achievement potential and intellectual efficiency include scales for Achievement by Conformance (*Ac*), Achievement by Independence (*Ai*), and Intellectual Efficiency (*Ie*). IV. Measures of intellectual and interest modes include scales for Psychological Mindedness (*Py*), Flexibility (*Fx*), and Femininity (*Fe*).

The extent to which these four groupings of the 18 scales are empirically justified by factor analysis has been evaluated critically in three separate factor studies of the CPI: (*a*) Crites, Bechtoldt, Goodstein, and Heilbrun (3); (*b*) Leton and Walter (8); and (*c*) Mitchell and Pierce-Jones (9).[2] Only the first of Gough's groupings appears as a well-defined factor in all three of these studies; the three other groupings have been called into question to a greater or lesser degree.

At the same time, a comparison of the factors derived from these studies shows a reassuring degree of factorial similarity although the populations studied differ in age, academic achievement, and client status. In all of these studies, however, the populations studied were relatively heterogeneous in intellectual ability. To assess the effect on factor definition of differences in intellectual ability in the population sampled, we undertook a factor analysis of three matrices of intercorrelations among the 18 scales of the CPI based on three groups with different levels of intellectual ability.

B. PROCEDURE

Three hundred male college freshmen and sophomores were formed into three groups of 100 each. (Their means and ranges on the ACE Psychological Examination are shown in Table 1.) These groups—a high, medium, and low—were essentially nonoverlapping in terms of ACE scores. ACE scores were obtained from college records. CPI scores were obtained by individual administration. The correlation matrices,[3] the factor analysis, and the factor rotation were computed on an IBM 7090 electronic computer. The principal-axis factoring method was used, with communalities estimated by the highest correlation of each scale with all other scales. Orthogonal Varimax was selected as the rotation criterion.

---

[2] Hereafter these three studies will be referred to respectively as Crites, Leton, and Mitchell.

[3] Correlation matrices are on file with the American Documentation Institute. Order Document No. 7430 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D.C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

TABLE 1
ACE[a] SAMPLE CHARACTERISTICS

Group	Range	Mean	*SD*	*N*
High (H)	133 – 183	152.10	11.94	100
Medium (M)	117 – 133	125.55	5.09	100
Low (L)	076 – 117	103.90	10.31	100

[a] All scores were converted to the 1952 ACE scale whenever other editions of the test had been administered. The 19 subjects with scores on the dividing lines between groups (117 and 133) were randomly assigned in order to achieve equal group size.

C. RESULTS, INTERPRETATION AND DISCUSSION

1. *Number of Significant Factors*

Table 2 shows the results of the application of several criteria for the termination of factoring. As can be seen from this table, these criteria offer considerable latitude to the researcher in deciding how many factors to rotate. Since mounting evidence based on the maximum-likelihood method for estimating factor weights indicates that use of rule-of-thumb criteria "tends to underestimate the number of factors that are statistically significant" (6, p. 378), the effects of rotating more factors than were generally indicated by the criteria in Table 2 were explored to assist in deciding on the number of factors justified.

Since previous factor analyses of the CPI have reported meaningful interpretation of four- and five-factor solutions, the factor solutions found

TABLE 2
NUMBER OF FACTORS INDICATED BY APPLICATION OF VARIOUS CRITERIA

	ACE group		
Criterion	High	Medium	Low
Humphries[a]	3	3	3
Guilford-Lacey[b]	5	4	5
Burt's Empirical[b]	3	2	3
McNemar[b]	3	4	4
Residuals ($SD \geqslant .026$)	5	6	7
100% original communality	4	4	4
Increase in communality $\geqslant 4\%$[c]	4	5	5

[a] See Fruchter (4, p. 80).
[b] See Thomson (11, pp. 121-123).
[c] See Harman (6, p. 363).

by rotation of five, six, and seven factors were compared for each ACE group separately. The ratio of the root-mean-square index to its value for "no agreement" was computed for judging similarity of factors (6, pp. 256-259). The results of these comparisons can be summarized[4] as follows: The first three factors are almost unchanged for *all* rotations in each ACE group and average 2.8 per cent dissimilarity. Assuming an arbitrary cutoff of 25 per cent dissimilarity as indicative of an unstable factor, the results indicate that five factors may be considered reasonably stable for all three ACE groups over a range of rotations with five to seven factors. On the basis of these results, plus the fact that the sixth factor showed no loading as high as .40 for any of the ACE groups when six factors were rotated, the five-factor solution was selected for the subsequent analyses and interpretation. Among the various criteria for the termination of factoring shown in Table 2, the Guilford-Lacey and the "increase in communality" criteria agree most closely with this judgment.

2. *Comparison of Factors Among Groups Differing in Intellectual Ability*[5]

In identifying and interpreting the factors obtained, comparisons were made of factors with similar patterns of loadings from each of the three ACE groups. The sizable correspondence in patterns of factor loadings over the three samples made it possible to group the results in terms of one set of five distinguishable factors (see Table 3).

The product moment correlation among the factor loadings reported in Table 4 served as an index of the degree of agreement[6] among loading patterns and was a primary basis for the factor grouping in Table 3. Each of the five factors was then interpreted from the standpoint of its pattern of loading in the three samples. Also included in Table 4 are intercorrelations with the factors reported by Mitchell, Leton, and Crites. Thus, both cross- and intra-study factor comparisons may be made.

a. Factor A: personal integrity and mental health. Factor A is a stable, general factor found in all three groups with highest loadings on CPI

---

[4] A table of the *rms* and *R* values obtained in these comparisons is on file with the ADI (see footnote 3).

[5] For ease of interpretation, results and discussion are presented jointly for each factor.

[6] The 18 scales enter as "individuals" in this correlation; however, since they are not entirely independent of one another due to overlap of their item composition, no assignment of levels of statistical significance is made to this index. Of course, such correlations should not be interpreted as correlations between factors in a common factor space. They serve only as an index of the extent to which factors from different samples have a similar pattern of factor loadings.

TABLE 3

FACTOR LOADINGS OF THE CALIFORNIA PSYCHOLOGICAL INVENTORY (CPI) SCALES FOR THE HIGH (H), MEDIUM (M), AND LOW (L) ABILITY GROUPS

CPI scale	Group	Factor					CPI scale	Group	Factor				
		A	B	C	D	E			A	B	C	D	E
Do	High	26	74	—14	08	—03	*To*	High	78	21	30	—01	11
	Medium	26	83	—06	05	—01		Medium	67	28	44	06	—03
	Low	27	81	—05	04	05		Low	50	26	69	—02	17
Cs	High	35	69	25	—09	02	*Gi*	High	80	22	—07	08	—19
	Medium	33	66	32	04	—07		Medium	83	16	—07	00	—33
	Low	29	63	39	—18	—04		Low	85	22	10	—08	—12
Sy	High	21	84	—01	—09	23	*Cm*	High	06	13	—17	06	51
	Medium	29	79	00	—16	03		Medium	—04	05	—15	—04	58
	Low	20	74	08	—34	17		Low	00	00	—07	—05	52
Sp	High	08	72	27	—42	06	*Ac*	High	77	26	—01	23	23
	Medium	03	64	39	—37	—12		Medium	76	35	03	12	—02
	Low	06	55	33	—48	—17		Low	62	42	28	00	29
Sa	High	—13	85	01	—07	09	*Ai*	High	50	04	62	18	—16
	Medium	—03	81	09	—18	02		Medium	46	19	66	21	—22
	Low	—05	79	05	—02	00		Low	27	00	79	18	—05
Wb	High	81	15	07	—25	22	*Ie*	High	61	42	32	—03	28
	Medium	77	26	12	—34	03		Medium	49	56	41	—02	14
	Low	64	22	37	—31	21		Low	34	42	62	—04	19
Re	High	65	18	—02	44	24	*Py*	High	58	07	18	—20	—14
	Medium	60	35	09	41	14		Medium	27	32	30	—02	—44
	Low	48	20	24	38	34		Low	34	23	47	—12	—25
So	High	53	08	—09	34	40	*Fx*	High	—01	06	74	—03	—15
	Medium	57	14	—12	—05	37		Medium	—16	06	70	05	—33
	Low	47	07	00	17	55		Low	—10	—10	60	—10	—41
Sc	High	87	—13	—01	13	05	*Fe*	High	02	—22	07	61	06
	Medium	91	—07	11	03	—15		Medium	—01	—21	10	53	—07
	Low	86	—03	25	05	10		Low	—02	—13	03	56	—02
Sum of squared loadings (S)								High	5.18	3.46	1.36	1.10	.85
								Medium	4.64	3.76	1.73	.83	1.02
								Low	3.48	3.17	2.69	1.03	1.20

Note: Decimal points omitted.

TABLE 4

CROSS-SAMPLE COMPARISONS: CORRELATIONS AMONG FACTOR LOADINGS

Factor	ACE group			Leton and Walter	Crites, *et al.*
	High	Medium	Low		
A: *Personal integrity and mental health*					
Mitchell & Pierce-Jones (I)	98	97	81	85	89
High ACE		96	94	88	85
Medium ACE			97	87	89
Low ACE				78	96
Leton & Walter (I)					63
B: *Social poise or extroversion*					
Mitchell & Pierce-Jones (II)	96	94	91	91	97
High ACE		96	96	92	92
Medium ACE			97	88	93
Low ACE				90	96
Leton & Walter (II)					89
C: *Capacity for flexible and independent thought and action*					
Mitchell & Pierce-Jones (IV)	89	92	91	78	93
High ACE		98	86	82	91
Medium ACE			91	85	94
Low ACE				75	95
Leton & Walter (IV)					81
D: *Femininity*					
Mitchell & Pierce-Jones (III)	72	46	65	81	64
High ACE		87	94	84	93
Medium ACE			89	62	86
Low ACE				76	93
Leton & Walter (III)					87
E: *Contented normativism*					
Mitchell & Pierce-Jones (III)	74	80	84	55	82
High ACE		93	90	84	91
Medium ACE			89	85	88
Low ACE				65	92
Leton & Walter (–V)					77

Note: Decimal points omitted. To help identify factors, the Roman numeral and capital letter factor tags employed in the original studies are included in parentheses. Reflected factors are indicated by a negative sign preceding the tag.

scales *Sc*, *Gi*, *Wb*, *Ae*, *To*, and *Re*, and sizable loadings on *So*, *Ai*, *Ie*, and *Py*. The generality of the factor is indicated not only by the large number of scales with high loading on this factor, but also by the large per cent of variance associated with it. High scorers on these scales are described by Gough (5) with such adjectives as deliberate, conscientious, self-approving, warm, cooperative, productive and organized, etc. Low scorers are seen as self-centered, shrewd, wary, insecure, constricted in thought and action, easily disorganized under pressure to conform, etc.

Factor A, which we have labelled "Personal Integrity and Mental Health," is highly similar for all three ability groups. This is indicated by the high intercorrelations (average .96) among the A_H,[7] A_M and A_L factors (see Table 4) and by the fact that the first six scales noted above are the primary defining scales in terms of factor loading size for all three ability levels.

Alongside of this pattern of factor similarity, there is an interesting change in the generality or breadth of factor A across ability levels. The total factor variance associated with factor A drops such that the sum of squared loadings for factor $A_H > A_M > A_L$ with A_L falling to 2/3 of A_H.

Cross-study intercorrelations presented in Table 4 indicate that factor A is highly similar to Mitchell's "Adjustment by Social Conformity" factor, Leton's "Mental Health and Personal Efficiency," and Crites' first factor. In the Crites' study each of their five factors is defined in terms of a single reference scale, in this instance the *Gi* scale.

All but three of the scales with loadings greater than .4 in the high ability group show successive decreases in the medium and low ability level groups. For the *Ai*, *Ie*, and *Py* scales loadings fall below .4 in the low ability group. This general drop in loadings on factor A suggests that this general factor is more restricted in scope at the lower intellectual levels. Since the low ability subjects are sampled from a college student population, it would be interesting to determine if this relationship is evident in even lower ability groups. Scales which show sizable loading on this factor for the brighter student show larger loadings on other factors for the less bright student. Thus, *Ie*, *To*, and *Ai*, the scales with the three highest loadings on factor C_L, all have sizable loadings ($\geqslant .5$) on factor A_H. However, loadings for the same scales on factor A_L average .26 less than for A_H. It appears that for the low ability group, scales involving capacity for independent thought and action (see discussion of factor C) are more associated with factor C_L than with the personal integrity factor.

These findings suggest that factor A is more broadly defined in the high ability group and includes aspects of intellectual efficiency and achievement. As group ability decreases, the behavioral scope of the factor narrows. In the low ability group intellectual efficiency and achievement are primarily associated with another factor. One may speculate that this relationship indicates how the poorer student may maintain personal integrity in the face of modest intellectual achievement. For him, scales descriptive of intellectual achievement tend to be more separate from the core of scales descriptive of his general mental health.

---

[7] Because of the large number of cross-sample comparisons among factors, it is convenient to use as a common reference system the letters A through E for factor identification and, when appropriate, to indicate by subscripts the particular ability sample from which the factor was derived.

b. Factor B: social poise or extroversion. Tables 3 and 4 show a high level of similarity in the definition of factor B for all three groups in terms of the high intercorrelations (average .96) among B_H, B_M and B_L and in terms of high loadings on *Sa*, *Sy*, *Do*, *Sp*, and *Cs*. These are five of the six scales that Gough (5) grouped together as "Measures of Poise, Ascendancy and Self-Assurance." It is also clear the Gough's inclusion of the *Wb* scale in this class is apparently without factor analytic support. Table 4 also shows the very high interstudy consistency of this factor. Mitchell names this factor "Social Poise or Extroversion"; Leton labels it "Social Confidence and Drive"; and Crites defines it in terms of the *Do* scale—"adjustment by control of external reality." Adjectives describing high scorers on these scales are outgoing, fluent in thought, aggressive, persuasive, confident; low scorers are seen as conservative, quiet, submissive, retiring, and awkward. It may be noted that, for our male college students, the *Ie* scale shows loadings greater than .40 in all three ACE groups. Since the loading on this scale falls *below* this criterional value for the Mitchell population of teacher candidates (largely female), for Leton's high school ninth graders, and for Crites' mixed-sex groups, we are left to speculate about the relevance of the finding. Whatever the basis, intellectual efficiency in their groups is not associated with the social poise or extroversion factor, whereas for our groups of college males, intellectual efficiency enters into the definition of this factor. Unlike factor A, no major differences in interpretation of this factor are associated with differences in intellectual level.

c. Factor C: capacity for flexible and independent thought and action. This factor pattern is identifiable in all three groups in terms of the high intercorrelations among C_H, C_M and C_L (average .92) and by the fact that the same four scales are those with the highest loading for each ability level. For the three groups, the highest average loadings on factor C are for scales *Ai* (.69), *Fx* (.68), *To* (.48). and *Ie* (.45). The items with similar content on the four common scales suggest that positive scores indicate a tolerance for difference and change, relative freedom from prejudicial and dogmatic attitudes, and a capacity for efficient, resourceful, and goal-directed thought and action.

As with factor A, factor C shows an interesting shift in behavioral scope across ability levels. The total factor variance associated with C_L is approximately twice that of C_H. With decreasing group ability level, the *Ai*, *To*, and *Ie* scales show higher loadings. It may be recalled that these scales show the opposite trend on factor A. These findings indicate that behavior measured by the scales which define factor C is only minimally associated with factor A for the low ability group, whereas factor A, for the high group, largely determines such behavior.

Table 4 again shows the relatively high cross-study similarity in factor loading patterns. Crites labelled their factor "Flexibility." The Mitchell factor is slightly broader, including sizable loadings on scales *Cs* (.50) and *Sp* (.51) as well. Mitchell labelled their factor "Capacity for Independent Thought and Action." Leton have labelled theirs "Social Dependence and Personal Independence." Leton's factor shows lower loadings on the four defining scales, although its pattern of loadings correlates sizably with the other C factors. Since the Leton study is based on a ninth grade high school sample, this difference suggests the interesting hypothesis that the capacity for independent thought differentiates as a distinct factor with greater age and maturity.

d. Factor D: femininity. Factor D is most clearly distinguished from the other factors by the ***Fe*** scale which has its only high loading on this factor (average .57). The other two scales with moderate loadings are *Sp* (average —.42) and *Re* (average .41); each of these scales has higher loadings on other factors. Adjectives which are common to these three scales and which characterize high scorers on this factor are patient, gentle, conscientious and dependable, moderate, self-restrained, respectful and accepting of others, and alert to ethical and moral issues. Low scorers on this factor are seen as outgoing, expressive, active and vigorous, robust, masculine, undercontrolled and impulsive in behavior, impatient with delay and indecision, and influenced by personal bias, spite and dogmatism.

This factor was identified as a feminine-masculine dimension in both the Crites and the Leton studies. Leton's interpretation derived additional support from the high loading (.72) of sex membership on this factor in their study.

Mitchell's factor III is more general than the other D factors. Its lower relationship to the other D factors appears to be due to their failure to extract a fifth factor. It this be the case, it is conjectured that the Mitchell's factor III would have split into two factors, more closely comparable to the D and E factors of our study. In support of this hypothesis, a linear combination of the scale loadings on factors D and E produces a multiple correlation with Mitchell's III factor of the order of .95 for each of the three ability groups. Because of its relationship to factors D and E, Mitchell's III is intercorrelated with factors D and E in Table 3 to facilitate comparisons.

e. Factor E: contented normativism. There are only two scales with sizable loadings for all three groups on factor E. These are the *Cm* scale (average .54) which has its only high loading on this factor and the *So* scale (average .44) which has higher loadings on factor A. In addition, —*Py*, —*Fx* and *Re* contribute in a borderline fashion to factor E's definition but each appears in only one of the three ACE groups. Using scale *Cm* as a primary point of departure for characterizing this factor, there is a core of attributes which define it. High scorers on this factor are described as conscientious, reliable, sincere, conforming and conventional, and deferential to authority, custom, and tradition. Low scorers are seen as deceitful, changeable, undependable, cynical, egotistic, highly concerned with personal pleasures, and rebellious toward rules, restrictions, and constraints. Thus, factor E at its "high" end seems to reflect an amiable acceptance of things as they are, a satisfaction with the socially-defined rules of living. If "super-ego strength" is used to characterize this factor, as Mitchell labels their factor III, then we must conclude that the super ego being described is of the "other-centered" variety (10).

Table 4 indicates factor E is most similar to that obtained in the Crites' study. Their description of this factor as a "tendency to make 'modal' responses" is consistent with our own. For the reflected Leton factor (—V), the unusually high negative loading (—.60) on the *Gi* scale and the slightly lower loading on the *So* scale suggest this factor shifts in meaning from that for the three ACE groups. For their ninth grade population, Leton defines their factor (the low end of ours) in terms of attempting to "look good" and impress others. It is not apparent how to reconcile meaningfully the differences between the E factors for the ACE groups and that reported by Leton.

D. SUMMARY AND CONCLUSIONS

The purpose of this study was to explore the factor stability of a multiscale personality test by examining the effects of differences in intellectual ability in the population sampled. Eighteen scales of the California Psychological Inventory were factored and rotated to the orthogonal Varimax criterion for three male college groups, each consisting of 100 students of high, medium, and low ability levels. The factors obtained were compared for stability among these three groups and with those reported in three previous studies.

For all three ability groups, whether five, six, or seven factors were rotated, the first five factors remained remarkably invariant for each of these rotations. Thus, the extraction and rotation of additional factors does not obfuscate the basic factors over the range explored. This finding provides some empirically-based reassurance in the face of contradictory criteria for deciding when to stop factoring.

For all three ability groups, five factors were identified and described:

A. Personal Integrity and Mental Health.
B. Social Poise or Extroversion.
C. Capacity for Flexible and Independent Thought and Action.
D. Femininity.
E. Contented Normativism.

Factorial stability across different populations for the CPI was indicated by identification of all five factors in all three ability groups. In addition, factors A, B, and C were identified in three other studies and factors D and E in two of the others where five factors were extracted.

Within this larger context of factorial invariance, there were also interesting shifts in factor characterization associated with ability level. In the high ability group as compared to the low, factor A, the personal integrity factor, tended to be more general and accounted for more of the total variance. Scales of intellectual efficiency and achievement, which showed sizable loading on the A factor in the high ability group, tended to be primarily associated with the C factor in the low group. It is speculated that the reduced loading of these scales on the adjustment factor for students of low ability may be a functional response to modest achievement.

The fact that test scales may differ in factor composition for different samples indicates the importance of determining the factorial stability or instability of personality scales for relevant subpopulations before employing them for individual counseling. With increased availability of computer programs for calculation of individual factor score estimates, the establishment of subpopulation factor patterns may prove to be a powerful adjunct to the tools of the test counselor.

REFERENCES

1. AHMAVAARA, Y. Transformation analysis of factorial data. *Ann. Acad. Sci. Fenn., Ser. B*, 1954, **88**, No. 2, 1-150.
2. COHEN, J. Factors underlying Wechsler-Bellevue performance of three neuropsychiatric groups. *J. Abn. & Soc. Psychol.*, 1952, **47**, 359-365.
3. CRITES, J. O., BECHTOLDT, H. P., GOODSTEIN, L. D., & HEILBRUN, A. B., JR. A factor analysis of the California Psychological Inventory. *J. Appl. Psychol.*, 1961, **45**, 408-414.
4. FRUCHTER, B. Introduction to Factor Analysis. New York: Van Nostrand, 1954.

5. GOUGH, H. G. Manual for the California Psychological Inventory. Palo Alto, Calif.: Consulting Psychologists Press, 1957.
6. HARMAN, H. H. Modern Factor Analysis. Chicago: Univ. Chicago Press, 1960.
7. JONES, L. V. A factor analysis of the Stanford-Binet at four age levels. *Psychometrika*, 1949, **14**, 299-331.
8. LETON, D. A., & WALTER, S. A Factor Analysis of the California Psychological Inventory and Minnesota Counseling Inventory. Santa Monica, Calif.: System Devel. Corp., 1961, SP-204.
9. MITCHELL, J. V., & PIERCE-JONES, J. A factor analysis of Gough's California Psychological Inventory. *J. Consult. Psychol.*, 1960, **24**, 453-456.
10. RIESMAN, D. The Lonely Crowd. New Haven: Yale Univ. Press, 1950.
11. THOMSON, G. H. The Factorial Analysis of Human Ability. Boston: Houghton-Mifflin, 1956.
12. WHEELER, W. M., LITTLE, K. B., & LEHNER, G. F. J. The internal structure of the MMPI. *J. Consult. Psychol.*, 1951, **15**, 134-141.

Center for Research in System Sciences
System Development Corporation
2500 Colorado Avenue
Santa Monica, California

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 457-468.

A METHOD FOR DETECTING SUBGROUPS IN A POPULATION AND SPECIFYING THEIR MEMBERSHIP* [1]

Department of Psychology, University of California

J. A. GENGERELLI

A. PROBLEM

It is sometimes desirable to determine whether a population of persons may be considered as constituted of a number of subsets or whether it is, so to speak, homogeneous. An illustrative problem is that of nosological classes: is the population of the mentally ill made up of several distinct subsets of persons, or must we look upon it as a group whose members are distinguishable in the same sense that one Shetland pony is distinguishable from another? A second illustrative problem is that of "personality types." Here the analogous question is: do types exist or do persons merely vary continuously and at random from one another in various dimensions?

Published work bearing on this topic has its center of gravity in determining whether two or more already isolated groups are, in fact, distinct. Best known in this area is the material on discriminant functions (4, 5, 9), and the publications of Hotelling (8), Bose (2, 3), and Rao (12). The concept of latent structure (1, 7) approaches the above problem more closely, while the work by McQuitty (10, 11) addresses itself rather directly to it.

B. SUBSETS AS CLUSTERS IN TEST SPACE

If there is a large population of N persons measured on but one trait, say, stature, the problem of subsets assumes its simplest form. Here an examination of the distribution of values gives us a clue; clearly, if the distribution is bimodal or multimodal, we suspect that more than one class constitutes the total population. Serious difficulties arise, however, when each member of the population N is described by a number of variables. If, for example, each individual is measured on 20 variables, with varying degrees of correlation existing between the variables, the ordinary statistical procedures become inadequate.

The complete solution to the problem involves answers to two separate

---

* Received in the Editorial Office on January 21, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] Supported in part by the Holter Research Foundation, Helena, Montana.

questions: the first, already mentioned, concerns the existence of subsets in the population. The second concerns their membership, that is, the specification as to which of the several subsets each member of the population belongs. We shall address ourselves to these questions in turn.

Consider a population of individuals of size *N* each measured on *m* variables, with correlations between the variables arbitrary but specified. Let any given individual's *m* scores be specified as a single point in the "test" space, that is to say, the position of that point in the space is such that its projection on each of the *m* axes corresponds to the value of the person's score on the axis. For simplicity, each score on a test axis is expressed in standard form. The performances of the *N* persons on the *m* tests would thus be represented as a swarm of points, *N* in number, in the test space. To aid the intuition, let *m* be 3. The set of *N* points would then be distributed in an ordinary 3-dimensional space. Clearly, the set of points could be imagined to assume many arrangements. They might form a dense homogeneous mass; or they might be distributed as, say, three dense masses, with "empty space" separating them mutually. These two possibilities are given because they give us the all-important clue to the problem's solution. In the first instance, it is clearly idle to talk of the existence of subsets in the population *N*. The population of points forms a single "community." In the second case, matters are quite different, for here we have three distinct communities which are well separated one from another.

The situation is analogous to that obtaining for human communities. If there are two dense clusters of buildings separated by much empty space, we have no difficulty in perceiving the existence of two villages; whereas if a village by one name coalesces with a village by another name, we feel that the separation is artificial and that there exist not two entities, but one.

This will be our criterion for the existence of subsets in a population. If the population of *N* points in the *m* space is distributed into clusters with "empty space" between them, subsets will be asserted to exist.

The fundamental problem presents itself as to how much empty space must separate two clusters if we are to recognize the clusters as separate.

We shall define a subset *as an aggregate of points in the test space such that the distance between any two points in the set is less than the distance between any point in the set and any point not in it.* As an illustration, imagine *N* points distributed in 3-dimensional space as 2 spheres, A and B, not necessarily of the same diameter. Then if the point in sphere A which is nearest to sphere B is separated from it by a distance greater than that separating any two points in any one of the spheres, we decide that the population of N points is constituted of two subsets. Otherwise, not. The matter may be described differently by saying that two subsets exist only if the two spheres are separated by a distance greater than the diameter of the larger sphere.

It may be objected that this criterion of separation is too severe; that we should accept a smaller distance as sufficient. However, it should be borne in mind in the context of the conception of individuals as points in test space that the distance between individuals in this space is an index of their similarity. If two individuals in the same subset are less similar than two individuals from different subsets, the necessity for the concept of subset becomes mitigated; indeed, the very meaning of the concept is rendered questionable.

C. Method for Detecting the Existence of Subsets

It is desired to devise a completely general method for detecting the existence of subsets among N individuals when each person has been tested on m variables. The investigator begins with m columns of scores, with N scores in each column. These are the data which must provide an answer to the question: do subsets exist? If so, how many? Who are the members of each subset?

Consider two points P_1 and P_2 in the m-space. Each point is determined by m co-ordinates, viz., the scores in the m tests taken. We wish to determine the distance separating the two points in the test space. Were the axes of the space orthogonal, this could be achieved by means of the pythagorean theorem in its usual form. That is,

$$S_{12}^2 = (z_{1a} - z_{2a})^2 + (z_{1b} - z_{2b})^2 + \dots + (z_{1m} - z_{2m})^2 \tag{1}$$

Where S is the distance between the two points (persons) indicated in the subscripts; z represents a standard score, the first subscript identifying the person (point) and the second identifying the test (dimension). Unfortunately, psychological variables are correlated, hence the axes of the space are not orthogonal but oblique, the cosine of the angle separating any pair of dimensions representing the Pearson correlation between the corresponding tests. Formula (1) then becomes equation

$$\begin{aligned} S_{12}^2 = {} & \cos a, a (z_{1a} - z_{2a})(z_{1a} - z_{2a}) + \dots + \cos a, m (z_{1a} - z_{2a})(z_{1m} - z_{2m}) \\ & + \cos b, a (z_{1b} - z_{2b})(z_{1a} - z_{2a}) + \dots + \cos b, m (z_{1b} - z_{2b})(z_{1m} - z_{2m}) \\ & \vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ & + \cos m, a (z_{1m} - z_{2m})(z_{1a} - z_{2a}) + \dots + \cos m, m (z_{1m} - z_{2m})(z_{1m} - z_{2m}) \end{aligned} \tag{2}$$

This complex expression gives us the square of the distance between two points when the axes of the space are not mutually perpendicular.

Equation (2) can be expressed more concisely as,

$$S_{12}{}^2 = \sum_{j,k} r_{jk}\,(z_{1j} - z_{2j})\,(z_{1k} - z_{2k}) \tag{3}$$

Where r_{jk} is the Pearson correlation coefficient between tests j and k and represents the cosine of the angle between the dimensions j and k in the test space. The summation is to be taken for all possible permutations of the tests from a to m, inclusive, as can be seen in equation (2).

The quantity $S_{12}{}^2$ represents the square of the "distance" between persons 1 and 2 in the test space. We can extract the square root if we wish, although for the present purpose this is not strictly necessary. Clearly, we must repeat the above operations for each possible pair of persons. If there are N persons in the population considered, then there will be $\dfrac{(N)\ (N-1)}{2}$ such quantities. This represents a formidable arithmetical chore, but it is well within the routine capacities of high speed computers.

When these $\dfrac{(N)\ (N-1)}{2}$ distances have been computed, we will have a set of distances which, when distributed, will range from a smallest number (the distance between the two nearest points in the test space) to a largest (the distance between those two points in the test space farthest away from each other). It is the analysis of this distribution of distances which will determine whether subsets do or do not exist in the population N.

Consider a space of $m = 3$ (for simplicity) in which the population of N points is distributed in the form of a sphere (again for simplicity). We are to calculate the distance between all possible pairs of points and plot the distribution. The nature of this distribution need not be ascertained; we know that the base line will extend from 0, in the event two points have identical coordinates, to a maximum value D, namely, the diameter of the sphere (in this example). There will be no instances of a distance equal to $(D + \epsilon)$, where ϵ is some positive number. Thus the distribution will very probably have but one mode *and will descend to zero on the right side of the mode.*

Consider, next, a 3-dimensional space in which the N points are arranged in the form of two spheres of uniform density, of diameters D_1 and D_2, respectively, with $D_2 > D_1$ and separated by a distance $(D_2 + \Delta)$ where Δ is a number greater than zero. The quantity $(D_2 + \Delta)$ represents the smallest space interval between a point in one sphere and a point in the other; in different language, the empty space separating the two spheres is greater than the diameter of the larger sphere. If we compute the distances between all points *within* the spheres and make a distribution of them, the distribution, although of undetermined form, will extend from some value not less than zero to the value D_2. Let us now proceed to measure the distance between all possible pairs of points which are in different spheres and add these values to the distribution already made. Clearly, none of these new values will be as small as D_2, the largest value in the first set; in fact the smallest value in the new set of values will be $(D_2 + \Delta)$. Thus all entries in the second set will lie to the right of the largest value in the first set and there will be a discontinuity in the distribution, as illustrated in Figure 1.

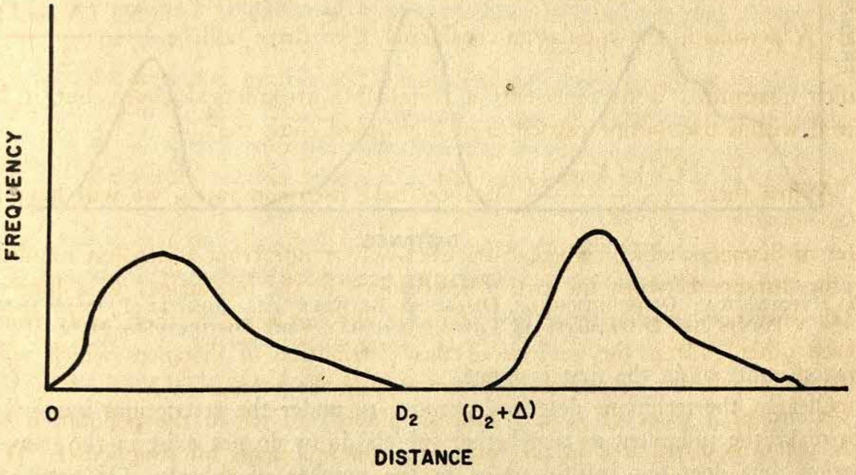


FIGURE 1
A HYPOTHETICAL DISTRIBUTION OF DISTANCES BETWEEN ALL POSSIBLE PAIRS WHEN N POINTS ARE DISTRIBUTED AS TWO VERY DISTINCT SUBGROUPS

Had the distribution of the N points been not in the form of two spheres but of two other shapes, however irregular, but such that the greatest distance between two points in one of the clusters were equal to D_1 and that between two points in the other equal to D_2; and, further, if the shortest distance between a point in one cluster and one in the other were equal to $(D_2 + \Delta)$, the resulting distribution of distances would be changed but there would be a discontinuity as before, beginning at D_2 and extending to $(D_2 + \Delta)$.

Were the N points distributed in three clusters rather than two and in a straight line such that one of the three clusters were interposed between the other two, with the shortest distance between any two adjacent clusters being greater than the distance between any two points in any cluster, then the resulting distribution of distances would have two discontinuities as shown in Figure 2. If, on the other hand, the three clusters while fulfilling the other conditions, were arranged in the test space not in a straight line but, for example, in the shape of an equilateral triangle, there would be but one discontinuity as in the first instance.

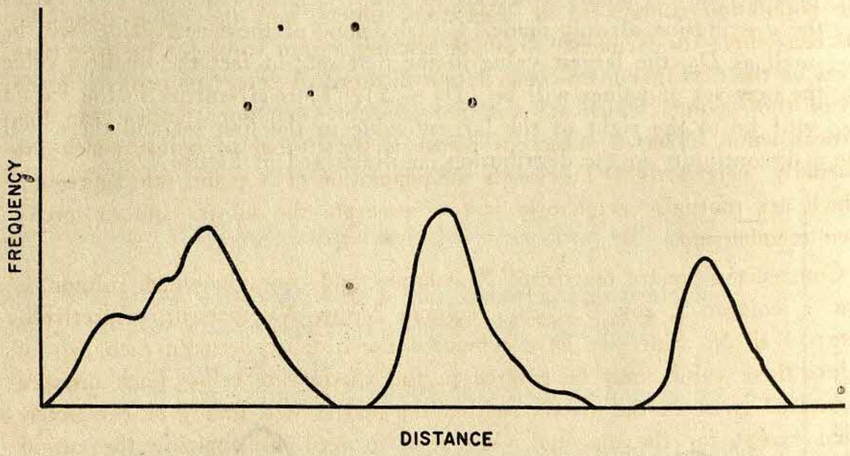


FIGURE 2
A Hypothetical Distribution of Distances Between All Possible Pairs When N Points Are Distributed as Three Mutually Very Distinct Subgroups

Clearly, the technique described permits us, under the restrictions imposed, to make the judgment as to whether subsets do or do not exist in the population, but does not inform us as to the number of subsets. Of greatest importance, is the fact that the procedure permits us to decide upon the "size" of the subsets when these exist. In particular, the first point of discontinuity in the distribution of distances gives us the size of the largest subset in the population. This furnishes us with the all-important criterion for determining whether two points are "neighborly" i.e., belong to the same subset. Thus if D_c is such a value, any two points at a distance equal to or less than D_c belong to the same subset while any two separated by a distance greater than D_c do not.

It should be borne in mind, of course, that if the interconstellation (or cluster) distances are equal to or smaller than the intracluster distances, the distribution of point distances will not show a discontinuity and hence choosing a criterion becomes more difficult. This is quite in keeping with our conception of a subset as an aggregate of points which are closer to one another than to points in any other set.

D. Subsets as Mutual Concurrences

The existence of a discontinuity in the distribution of distances tells us that the population consists of at least two subsets, and the value along the abscissa where the frequency ordinate assumes a zero value for the first time gives us the maximal infra-constellation distance. A subset of points is made up of those points which are not more distant from one another than this critical value. Thus a subset of points is constituted of points which are mutually "neighborly." To classify the population of N points into aggregates which are mutually neighborly is to enumerate the subsets and to specify their membership.

Construct a square matrix of N columns and rows, in which column 1, row 1, column 2, row 2 .... column N row N represent, respectively, points 1 to N. Since we have computed the distance between each pair of points these values may be entered in the appropriate cells. Each distance is entered twice in the matrix so that the matrix will finally be completely filled except for the diagonal. We then proceed, by applying the cut-off criterion already obtained, to ascertain whether a given pair of points are "neighbors" or "strangers" and make the appropriate entry in each cell. Let the letter *"N"* represent the former and *"S"* the latter.

The data are now in a form which is readily analyzable.

Sum the number of "N" cells in each vertical array of the square matrix. To each sum add the digit 1. The augmented sum for a given vertical array represents the maximal size of a subset of which that person (point) could be a member, provided all the other members are mutually "neighborly." The unaugmented sum represents the number of persons with whom the person in question is "neighborly." Consider any array, say, the first vertical array. Let the number of "N" entries be k; this indicates that X_1 is neighborly with k other individuals (points). By the definition of neighborliness, viz., a separation equal to or less than some critical distance D_c, the k points must therefore be in the same subset as X_1. The *"N"* entries in each of the vertical arrays corresponding to the k neighbors in the first vertical array will therefore show identical patterns. These columns may now be checked off since they have been identified as belonging to one subset.[2]

There will be (*N-k-1*) columns remaining whose membership is confined to one subset or may be divided among two or more subsets. The analysis is repeated with a vertical array representing a person not neighborly with X_1. When the work is finished a second subset and its membership will have been defined. If "unclaimed" vertical arrays remain, the process is repeated.

We have defined a subset as an aggregate of points (persons) such that any two points in the set will be separated by a smaller distance than any point in the set from any point not in it; or, in different language, we may say that subsets in a population exist when any infraset distance is smaller than any interset distance. As has been noted, this is a very severe criterion, and it is very doubtful that it would be satisfied by psychological data obtained from a single culture. Indeed, this criterion is not completely met even in astronomy. While, in general, the distance between two given aggregates of stars may be greater than the distance between any two stars in the same aggregate, the space between the two aggregates will not be completely empty. There will very often be a few isolated stars scattered throughout the intervening space. It is necessary, therefore, to investigate the effect of a similar state of affairs upon our method for discovering and isolating subsets.

Consider, once more, the example of two spheres, with the diameters such that $D_2 > D_1$, and the distance between the two nearest points in the two spheres equal to $(D_2 + \Delta)$. Imagine a point in the test space placed exactly half-way between the two spheres. What effect would this have on the distribution of distances between all possible pairs of the $N + 1$ points? Clearly, such alterations as will occur are affected by the distances existing between this intermediate point and each of the points in Sphere 1 and Sphere 2, respectively. The argument can best be followed by referring to Figure 1. Some of these new distances will fall into the aggregate of scores to the left of the discontinuity shown in the figure; others will very probably fall in the aggregate to the right. Still others, however, will almost certainly fall in the region of discontinuity; and if the number of points in each of the two spheres is very great, the effect will be to efface the discontinuity altogether and to replace it with a "transition bridge" of frequencies distinctly smaller in value than those on either side of it. The distribution would assume the appearance of a deep saddle.

---

[2] The argument here hinges upon the assumption that a point of sharp discontinuity has been found in the distribution of distances. In that case, if the distance between X_1 and each of the other k persons (points) is less than the critical distance D_c the k persons must be in the same constellation as X_1 because only in that way can S be less than D_c.

If, now, we imagine the space between the two spheres to be populated not by one point exactly midway but by a number of points randomly scattered throughout the intervening space, the effect will be to make the "saddle" less deep. If these points are few in number compared to those to be found in each of the spheres, however, the distribution will maintain a pronounced bimodal appearance. As the number of "intersphere" points increases the feature of bimodality becomes increasingly dissipated until, in the limit, when the density of points in the total test space is fairly uniform, it disappears altogether. The existence of subsets in a population is therefore associated with the feature of multimodality in the distribution of inter-point distances. When the ratio of intergroup distances to infragroup distances is less than 1 and very small, the distribution will be monomodal: as this ratio approaches the value 1, the distribution changes in the direction of multimodality, and when the ratio becomes greater than 1, the distribution develops a definite discontinuity, the size of the gap correlating positively with the size of the ratio.

Whether, in a given case, the distribution of interpoint distances is significantly nonmonomodal may be tested by conventional statistical means; it should be emphasized, however, that if such a test is necessary it is very unlikely that the enterprise will be worth while.

When the distribution of interpoint distances is obviosuly nonmonomodal but not disjunct, we are faced anew with the problem of determining the value of the critical distance which defines "neighborliness." Clearly, the cut off should lie somewhere along the bottom of the first "dip" in the distribution, i.e., somewhere along the x-axis in the region between the point where the negative derivative of the curve approaches zero and the point where it becomes positive again. The exact locus of the cut is of minor importance and, at any rate, is not deducible from general considerations. A safe practice is to make the cut at the point where the slope of curve becomes positive. This value on the distance-axis then becomes the critical value D_c, and a subset is then defined as the aggregate of points (persons) which are *mutually* no farther apart one from another in the test space than the distance D_c, i.e., no two persons in the subset are separated by a distance greater than D_c.

Applying the criterion, we construct the distances matrix, as described earlier, and enter the symbols "*N*" and "*S*" in the appropriate cells. As before, we sum the number of "*N*" entries in each vertical (horizontal) array and augment each sum by 1. The augmented sum in a given array represents the maximal size of the subset to which that point belongs, in the event all the other points are mutually neighborly. In the example considered earlier, it was inevitable that all points neighborly to a given point were neighborly to one another; in the present context, however, this need not be so by virtue of the fact that we have the existence of transition points which bridge the space between clusters. We must therefore apply the method of mutual concurrence analysis described in detail elsewhere (6).

Choose that column having the largest sum; if there are several, choose one at random. Let this be, say, column 3. Next choose a column having its sum equal to the one already chosen or nearest to it. Let this be, say, column 18. Refer now to the cell formed by the intersection of arrays 3 and 18. If the cell contains an "*N*," 18 is retained; if it contains an "*S*," 18 is rejected. If it is rejected, another column having the largest available sum is chosen; if it is retained we similarly proceed to the column with the largest available sum. Let us suppose that 18 is retained and that the next column is 12. We now refer to cells (3, 12) and (18, 12). If both have "*N*," element 12 is retained; if either or both have "S" it is rejected. As a final example, suppose the elements 3, 12, 18, 36, 37 have been retained, and the next element to be tested is 19. We now refer to cells (3, 19), (12, 19), (18, 19), (36, 19), (37, 19). If any of these cells contains the letter "*S*," column 19 is rejected; if all contain the letter "*N*," the column is accepted. We proceed in this manner until all the columns have been considered. When we have finished, the mutually neighborly points will constitute a subset. There will be a great many columns not included. We repeat the procedure, beginning with that column having the largest available sum and working through each of the remaining colums. When this second cycle is completed we will have the membership of the second subset. If there are columns which have not been included in either the first or second subset, a third cycle is initiated and carried through.

When no further columns remain, the analysis is complete: the population of *N* has been divided into the various subsets and the membership in each indicated. A synthetic problem illustrating the procedure is shown in Table 1.

As the analysis proceeds, the membership of the later subsets becomes smaller and smaller. The members of these tail-end subsets are, in fact, the "transition points" in the test space.

TABLE 1

Synthetic problem involving a "population" of 20, divided into two subsets constituted of the points (persons) 1 - 4 - 5 - 9 - 10 - 12 - 14 - 16 - 17 - 18 - 19 - 20 and 2 - 3 - 6 - 7 - 8 - 11 - 13 - 15, respectively. *N* and *S* represent "neighbors" and "strangers," respectively. *N*'s represent distances equal to or less than some critical value D_c; *S*'s represent distances greater than D_c. If random values conforming to this criterion were inserted in the matrix and a frequency distribution made, a bimodal curve similar to that shown in Figure 1 would be obtained.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1		S	S	N	N	S	S	S	N	N	S	N	S	N	S	N	N	N	N	N
2	S		N	S	S	N	N	N	S	S	N	S	N	S	N	S	S	S	S	S
3	S	N		S	S	N	N	N	S	S	N	S	N	S	N	S	S	S	S	S
4	N	S	S		N	S	S	S	N	N	S	N	S	N	S	N	N	N	N	N
5	N	S	S	N		S	S	S	N	N	S	N	S	N	S	N	N	N	N	N
6	S	N	N	S	S		N	N	S	S	N	S	N	S	N	S	S	S	S	S
7	S	N	N	S	S	N		N	S	S	N	S	N	S	N	S	S	S	S	S
8	S	N	N	S	S	N	N		S	S	N	S	N	S	N	S	S	S	S	S
9	N	S	S	N	N	S	S	S		N	S	N	S	N	S	N	N	N	N	N
10	N	S	S	N	N	S	S	S	N		S	N	S	N	S	N	N	N	N	N
11	S	N	N	S	S	N	N	N	S	S		S	N	S	N	S	S	S	S	S
12	N	S	S	N	N	S	S	S	N	N	S		S	N	S	N	N	N	N	N
13	S	N	N	S	S	N	N	N	S	S	N	S		S	N	S	S	S	S	S
14	N	S	S	N	N	S	S	S	N	N	S	N	S		S	N	N	N	N	N
15	S	N	N	S	S	N	N	N	S	S	N	S	N	S		S	S	S	S	S
16	N	S	S	N	N	S	S	S	N	N	S	N	S	N	S		N	N	N	N
17	N	S	S	N	N	S	S	S	N	N	S	N	S	N	S	N		N	N	N
18	N	S	S	N	N	S	S	S	N	N	S	N	S	N	S	N	N		N	N
19	N	S	S	N	N	S	S	S	N	N	S	N	S	N	S	N	N	N		N
20	N	S	S	N	N	S	S	S	N	N	S	N	S	N	S	N	N	N	N	
ΣN	11	7	7	11	11	7	7	7	11	11	7	11	7	11	7	11	11	11	11	11
ΣN+1	12	8	8	12	12	8	8	8	12	12	8	12	8	12	8	12	12	12	12	12

E. SUMMARY

A method is described whereby a population of *N* individuals, tested on *m* variables, may be examined for the existence of subgroups. The method involves the interpretation of the *m* scores for a given individual as a point in *m*-space and the population as a set of *N* points in the space. Distances between all possible pairs of individuals are computed and a distribution made. A criterion is developed which, applied to the distribution of distances, yields a judgment as to whether subgroups exist in the population. An application of the method of mutual concurrences is described for determining the membership of each subgroup.

REFERENCES

1. ANDERSON, T. W. On estimation of parameters in latent structure analysis. *Psychometrika*, 1954, **19**, 1-10.
2. BOSE, R. C., & ROY, S. N. The distribution of the studentized D^2 statistic. *Sankyā*, 1938-40, **4**, 19-38.

3. ———. The use and distribution of the studentized D^2 statistic when the variances and covariances are based on k samples. *Sankyā,* 1938-1940, **4**, 535-542.
4. FISHER, R. A. The use of multiple measurements in taxonomic problems. *Ann. Eugen.*, 1936, **7**, 179.
5. ———. The statistical utilization of multiple measurements. *Ann. Eugen.*, 1938, **8**, 376.
6. GENGERELLI, J. A. The analysis of mutual concurrences. *J. of Psychol.*, 1961, **52**, 253-263.
7. GREEN, B. A general solution for the latent class model of latent structure analysis. *Psychometrika*, 1951, **16**, 151-166.
8. HOTELLING, H. The generalization of Student's ratio. *Ann. Math. Statist.*, 1931, **2**, 360-368.
9. MCNEMAR, Q. Psychological Statistics: 3rd Ed. New York: John Wiley & Sons, 1962.
10. MCQUITTY, L. L. Elementary linkage analysis for isolating orthogonal and oblique types and typal relevancies. *Educ. Psychol. Meas.*, 1957, **17**, 207-229.
11. ———. Typal analysis. *Educ. Psychol. Meas.*, 1961, **21**, 677-696.
12. RAO, C. R. Advanced Statistical Methods in Biometrics Research. New York: John Wiley & Sons, 1952.

Department of Psychology
University of California
Los Angeles 24, California

Published as a separate and in *The Journal of Psychology,* 1963, **55**, 469-475.

ANAGRAM SOLUTION TIMES: A FUNCTION OF WORD LENGTH AND LETTER POSITION VARIABLES*

University of Southern California and New York University

M. S. MAYZNER AND M. E. TRESSELT

A. PURPOSE

In previous work (3, 4, 5) the effects of certain variables on anagram solution times were examined within the framework of an S-R mediation model of problem solving behavior (1, 2). It was suggested than an anagram constitutes a problem whose stimulus properties evoke a variety of implicit responses in *S,* and the four variables under study, namely, anagram letter order, word frequency, anagram transition probability, and word transition probability, were related to the way in which these implicit responses are produced.

B. METHOD

This study involves a further analysis of the data collected in the earlier work (3, 4, 5), based on new tables of digram and single-letter frequency counts broken down by word length (WL) and letter position (LP) combinations (7). The present analysis takes as its point of departure the previous finding (5) that anagrams whose word solutions have high summed digram frequency totals are solved significantly faster than anagrams whose word solutions have low summed digram frequency totals. In the earlier study (5), 18 words were employed, nine words having low summed digram frequency totals, yielding a median solution time of 28.00 seconds, and nine words having high summed digram frequency totals, yielding a median solution time of 14.40 seconds. However, if the rank-difference correlation between median solution times and summed digram frequency totals is calculated, its value is only .04. Thus, while the effect is appreciable (28.00 seconds versus 14.40 seconds) when comparing median solution times based on groups of words having either low or high summed digram frequency totals, this variable as presently measured is not sufficiently strong to yield an appreciable rank-difference correlation. In calculating this rank-difference correlation of .04, Underwood and Schulz's norms (8) were employed in obtaining the summed digram frequency values. With the

---

* Received in the Editorial Office on January 21, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

frequency norms now available (7), which take into account the effects of WL and LP combinations on digram frequency values, a finer analysis can be made of the previous data.

Before presenting this analysis, the table giving the digram frequency counts on which the analysis is based, should be described. Following Underwood and Schulz's procedure (8), 100 samples of 200 words each were collected from a wide variety of newspapers, magazines, and fiction and nonfiction books. For each source selected, a starting place was chosen at random and proceeding forward from this point all three-, four-, five-, six-, and seven-letter words were recorded until a total of 200 words had been selected. This procedure was duplicated 100 times, each time for a different source. A grand total of 20,000 words was thus obtained. This sample of 20,000 words consists of 6,807 three-letter words, 5,456 four-letter words, 3,422 five-letter words, 2,264 six-letter words, and 2,051 seven-letter words. Digram and single-letter frequency counts were then made for all WL and LP combinations. Only the digram counts from these tables are used in the present study.

For each word used in the several anagram studies under consideration three digram frequency values were obtained. The method employed in obtaining these values may best be understood by referring to Table 1. In Table 1, the word "sugar" is shown broken down into its four sequential digrams "su," "ug," "ga," and "ar," and for each digram three digram frequency values were obtained from the table of digram frequency norms (7). For example, the digram "su" occurs seven times in the first and second position of a five-letter word, it occurs 17 times in five-letter words independent of letter position, and it occurs 175 times in the total sample of 20,000 words independent of word length or letter position. This procedure yields three summed digram frequency values for the word "sugar," i.e., a value of 35 by WL and LP, a value of 238 by WL, and a value of 1219 by Total, the Total

TABLE 1
DIGRAM FREQUENCY COUNTS FOR THE WORD "SUGAR"

Digrams	Frequency counts		
	WL & LP	WL	Total
su	7	17	175
ug	4	14	148
ga	11	33	94
ar	13	174	802
Sum	35	238	1219

value being analogous to the value obtained when employing Underwood and Schulz's norms (8).

C. RESULTS

Table 2 presents all of the words employed in the three anagram experiments—Experiment 1 (5), Experiment 2 (3), and Experiment 3 (4)—under consideration, the three digram frequency values obtained by employing the method described in Table 1, and the solution times to the anagram form of the word. Since in Experiments 2 and 3, the words employed varied in their Thorndike-Lorge frequency values, they are grouped separately by frequency groups (i.e., Group VF and F, which are combined since they are "AA" and "A" words; Group I, which are words occurring at least once but not twice per million; and Group VI, which are words occurring less than once per million but more than once per four million).

Table 3 presents the rank-difference correlations (Rho) between word digram frequency counts and anagram solution times (with the lowest frequency count and the slowest solution time both ranked "one") for the several conditions of Experiments 1, 2, and 3 given in Table 2. In Table 3, 12 conditions for calculating Rho are provided and for 10 of these conditions we find a systematic increase in the value of Rho going from Total to WL to WL and LP. The probability of finding such a pattern, as calculated by the binomial expansion[1] $P(r)(n\text{-}r) = nCr\, p^r q^{n-r}$, when $r = 10$, $n = 12$, $p = 1/6$, and $q = 5/6$, is approximately once in a million occurrences.

Looking specifically, for example, at the value of Rho in Experiment 1 under the Total column, in Table 3, we find a value of —.01, which is essentially equivalent to the value of .04 obtained when employing Underwood and Schulz's norms. However, when WL norms are employed Rho increases to a value of .15 and when WL and LP norms are employed Rho further increases to a value of .41, and this type of systematic increase is found for 10 of the 12 conditions examined. Thus it would appear that as the digram frequency structure of the word is more precisely characterized by taking word length and letter position variations into account, a much stronger relationship emerges between word digram structure and solution time.

Since it has also been found in previous work (4) that the digram frequency count associated with the anagram rather than the word also affects solution time, digram frequency counts based on Total, WL, and WL and

---

[1] The authors would like to express their appreciation to Dr. Arnold Binder, who suggested the use of the binomial expansion to evaluate the probability of finding such a pattern of correlation values.

TABLE 2
WORD DIGRAM FREQUENCY COUNTS AND ANAGRAM SOLUTION TIMES FROM THREE DIFFERENT EXPERIMENTS

	Digram frequency counts						
Words	WL & LP	WL	Total	Solution times			
			Experiment 1				
sugar	35	238	1219	83.0			
fruit	44	136	900	15.0			
music	45	243	902	27.0			
judge	49	79	412	3.0			
value	49	183	843	46.0			
guide	67	173	757	7.0			
uncle	84	316	1075	72.0			
human	91	235	2132	15.0			
guard	142	248	1037	49.0			
month	168	661	5039	13.0			
voice	183	337	708	4.0			
paint	237	380	1952	13.0			
drink	250	363	1565	7.0			
power	303	529	2146	22.0			
youth	336	907	5720	18.0			
pound	363	556	2755	17.0			
water	443	733	3277	3.0			
house	452	698	2527	6.0			
			Experiments 2 and 3				
Groups VF & F[b]				E[a]	H	LT	HT
sugar	35	238	1219	10.5	91.5	182.5	88.5
labor	69	460	1527	5.0	210.0	159.5	201.0
giant	104	224	2072	7.5	154.0	29.0	240.0
model	112	188	1115	4.5	191.5	62.5	59.0
party	136	304	1255	14.0	24.0	51.0	65.0
train	167	390	1908	5.0	8.0	12.0	118.5
fault	168	195	553	7.0	36.0	56.0	136.5
chair	186	441	2170	10.0	12.5	8.5	57.0
clerk	208	607	2066	17.5	22.0	31.5	91.5
beach	352	520	1805	3.0	9.5	8.5	37.5
Group I							
baton	65	267	1850	9.5	194.0	85.0	240.0
cobra	79	160	742	50.5	240.0	240.0	240.0
patio	90	180	1267	22.0	125.0	71.0	218.0
jaunt	110	185	745	5.0	37.5	33.0	240.0
roach	118	351	1289	9.5	19.5	38.0	80.5
Group VI							
tango	78	391	2741	45.0	240.0	240.0	240.0
peony	81	165	1037	12.0	240.0	45.5	240.0
triad	110	187	900	5.0	103.0	61.5	240.0
groin	141	411	1800	31.0	35.0	25.5	240.0
ghoul	187	728	2137	19.5	33.5	75.5	240.0

[a] E, H, LT, and HT refer to the anagram form used, i.e., easy order, hard order, low digram frequency total, and high digram frequency total, respectively.

[b] Groups VF, F, I, and VI refer to Thorndike-Lorge word frequency values from very frequent to very infrequent.

TABLE 3

RANK-DIFFERENCE CORRELATIONS (RHO) BETWEEN WORD (W) AND ANAGRAM (A) DIGRAM FREQUENCY COUNTS AND ANAGRAM SOLUTION TIMES FROM THREE DIFFERENT EXPERIMENTS

	Digram frequency count conditions					
Experimental conditions	WL & LP		WL		Total	
Experiment 1	W	(A)	W	(A)	W	(A)
(18 words)	.41	(.21)	.15	(—.17)	—.01	(.02)
Experiments 2 and 3						
(10 Group VF & F words)						
E	.06	(.30)	—.13	(.31)	—.20	(.26)
H	.76	(.09)	.48	(.16)	.42	(—.05)
LT	.75	(—.03)	.42	(.06)	.73	(—.04)
HT	.48	(.36)	.20	(—.27)	—.01	(—.24)
(5 Group I words)						
E	.42	(.37)	.66	(.62)	.37	(.37)
H	.90	(—.40)	.70	(—.20)	.30	(—.30)
LT	.80	(.67)	.50	(.50)	.10	(.60)
HT	.70	(—.10)	.50	(—.10)	.40	(.50)
(5 Group VI words)						
E	.20	(.90)	—.50	(.50)	—.90	(.90)
H	.97	(.82)	.82	(.67)	.07	(.67)
LT	.30	(—.20)	—.20	(—.60)	—.60	(—.60)
HT	—[a]	—	—	—	—	—

[a] Solution times for all five words were 240.0 seconds, therefore no correlations are possible.

LP norms were calculated for all of the anagrams employed in Experiments 1, 2, and 3. Rank-difference correlations were then calculated (with the lowest frequency count and the fastest solution time both ranked "one") for the several conditions and are presented in Table 3 in parentheses immediately after the correlations based on the words. Here no systematic pattern is evident and apparently no relationship exists between a finer characterization of anagram digram structure and solution time.

The results reported in Table 3 provide additional support for the mediational model that is being employed. As previously indicated (3, 4, 5) the model assumes that an anagram constitutes a problem whose stimulus properties evoke a variety of implicit responses in *S*, and that at least some fraction of these implicit responses consist of two-letter groupings or digrams. Further, the production of these digrams reflect their position in *S*'s response hierarchy, which is a function of his exposure to the language over a period of several years. Thus, when *S* implicitly begins forming digrams with the letters given to him in a particular anagram problem, he tends first to produce digrams high in his response hierarchy (i.e., digrams with high frequencies of occurrence in the language and therefore high in his response hierarchy) and this is what is assumed to lead to faster solution times for those words which are made up of high frequency digrams. The specific findings of this study now suggest that not only does *S* produce high frequency digrams first, but that this implicit production process takes account of WL and LP characteristics of the problem as indicated by the higher correlations found when based on WL and LP norms. Additional support for the conception that *S* is sensitive to WL and LP combinations comes from another type of study (6) in which *S* was required to rank the frequency of occurrence of digrams and single letters as a function of WL and LP combinations and it was found that *S*s could make such estimates to a highly significant degree.

D. CONCLUSIONS

The overall results of the studies under discussion (3, 4, 5, 6) uniformly point to the critical importance of the WL and LP variable and suggest a further test of our mediational model. Namely, if *S*s are sensitive to WL and LP digram frequency norms, and the production of their implicit digram responses reflects the frequency-hierarchy of such norms, then anagrams whose word solutions require the placement of digrams in their most frequently occurring position in a word of a given length should be solved faster than those which require the placement of digrams in infrequently occurring positions. For example, if *S* is presented with the letters "eoprst" which when properly rearranged yields either poster or presto, he should require much more time to solve for presto than for poster, since the WL and LP digram frequencies associated with "pr," "re," "es," "st," and "to" are 22, 16, 4, 7, and 2, respectively, summing to 51, while the WL and LP digram frequencies associated with "po," "os," "st," "te," and "er" are 23, 8, 25, 104, and 193, respectively, summing to 353. If we examine the digram frequencies independent of WL and LP for these words and sum, a value of 2964 is found for presto and a value of 2995 is found for poster, which would lead to the prediction that both words are equally difficult. However, by taking the WL and LP variable into account, a marked difference in WL and LP digram frequencies is found (51 versus 353) leading to the prediction that presto should be much more difficult to solve for than poster. Since Thorndike-Lorge word frequency is related to anagram solution time, it is essential that both presto and poster have approximately the same word frequency, if differences in solution time are to be related to digram frequencies, and both do have approximately equal word frequencies, with presto occurring once per million and poster five times per million. Problems of the "presto-poster" type should provide then an additional and critical test of our mediational model and of the importance of the WL and LP variable for the model.

REFERENCES

1. KENDLER, H. H. Problems in problem solving research. In *Current Trends in Psychological Theory: A Bicentennial Program.* Pittsburgh: Univ. Pittsburgh Press, 1961.
2. KENDLER, H. H., & KENDLER, T. S. Vertical and horizontal processes in problem solving. *Psychol. Rev.,* 1962, **69**, 1-16.
3. MAYZNER, M. S., & TRESSELT, M. E. Anagram solution times: A function of letter order and word frequency. *J. Exper. Psychol.,* 1958, **56**, 376-379.
4. ———. Anagram solution times: A function of transition probabilities. *J. of Psychol.,* 1959, **47**, 117-125.
5. ———. Anagram solution times: A function of word transition probabilities. *J. Exper. Psychol.,* 1962, **63**, 510-513.
6. ———. The ranking of letter pairs and single letters to match digram and single-letter frequency counts. *J. Verb. Learning & Verb. Behav.,* 1962, **1**, 203-207.
7. ———. Tables of digram and single-letter frequency counts by word length and letter position for a sample of 20,000 words. Mimeographed copy, 1962.
8. UNDERWOOD, B. J., & SCHULZ, R. W. Meaningfulness and Verbal Learning. Philadelphia: Lippincott, 1960.

Department of Psychology
University of Southern California
University Park
Los Angeles 7, California

Published as a separate and in *The Journal of Psychology*, 1963, **55**, 477-485.

CLASSIFICATION TRANSFER WITH DISJUNCTIVE CONCEPTS AS A FUNCTION OF VERBAL TRAINING AND SET*[1]

Department of Psychology and Psychiatry, Catholic University

HANS G. FURTH

A. PROBLEM

Verbal behavior is widely used as a means of dealing experimentally with concept knowledge but seems to leave the critical investigator in doubt as to whether it satisfactorily reflects cognitive processes. Nonverbal methods have the advantage of avoiding the complex language-cognition relationship; on the other hand, they provide uncertain criteria of concept attainment. Moreover, a particular nonverbal task is usually ambiguous as to the determining factor in success or failure. Memory, rigidity, perceptual set, an indefinite number of chance factors influencing discovery of the relevant aspect, may to a greater or lesser degree influence behavior which is labeled concept attainment.

To obviate some of these difficulties the present author proposes the idea of "level of concept attainment" and attempts to test it by means of transfer. This notion is not new: conceptual knowledge is best manifested by the way it is utilized in a variety of new situations. For experimental purposes a person first is made familiar with a certain concept and is challenged to use the now familiar concept in a different situation. Ease and rapidity of utilization of the concept in the new problem situation should be in direct relation to the previous level of concept attainment. Thus, we can obtain an answer not merely to: "Was the concept attained?", but to the question: "How well was it attained?"

In an effort to discover concepts which were suitable for a nonverbal procedure and which could be employed meaningfully in a transfer task, concepts of the disjunctive type were employed. A disjunctive concept is defined as one which is expressed by "either . . . or" and requires the

---

* Received in the Editorial Office on January 22, 1963, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

[1] This study was supported in part by the United States Office of Vocational Rehabilitation research grant RD-704. The author wishes to thank Alice Lazar, Lawrence and Carolyn Brick for help in the collection of data. The cooperation of Dr. G. E. Detmold and Mr. R. M. Phillips and staff and students at Gallaudet College is gratefully acknowledged.

inclusion of two separate classes into one supra-ordinate class. A good example would be the class of persons permitted to use a college elevator as indicated by the sign: "For faculty members and disabled persons only." Apart from the obvious possibility of a disabled faculty member, such a sign is correctly understood as implying the either-or dichotomy of a disjunctive concept.

Investigators working in the area of concept formation have observed that disjunctive concepts are obtained with greater difficulty than other kinds of concepts. Hunt and Hovland (3) simulating concept attainment by computer process showed that college students will choose disjunctive concepts with significantly less frequency than relational or conjunctive concepts. Bruner, Goodnow, and Austin (1, Chap. 6) came to a similar conclusion with slightly different procedures. In contrast to other classes the formation of a disjunctive class is more likely to require what Piaget (8) calls the "formal operation of inclusion" and does not have concrete perceptual support to the same extent as other classifications.

This study proposed to familiarize *S*s first with the four components of a classification problem by way of attainment of one disjunctive and two separate conceptual categories. The one disjunctive category included two separate concepts. The level of concept attainment was tested by a subsequent shift in the disjunctive combination with the expectation that persons impoverished in verbal training would transfer with less ease and rapidity than a control group. In this connection it was reasonably assumed that persons, deaf from early childhood, have had considerably less training in and experience with verbal language than hearing persons.

To control for the possibility that different types of shift would result in a difference of performance four transfer groups were observed. In this way the influence of set on transfer with a disjunctive concept and interaction of set and level of concept attainment could be observed. Finally various theoretical models of transfer would be evaluated in light of the results on the shift tasks.

B. METHOD

1. *Subjects*

The hearing *S*s of this study were four different classes of first or second year college students at Catholic University, one in education and three from R.O.T.C. training. The four deaf groups were four classes at Gallaudet College. The four groups of hearing classes comprised 38, 39, 40 and 24 students, while the corresponding number for the deaf were 20, 25, 26 and

28. Gallaudet College draws its population from the best graduates of schools for the deaf throughout the country and thus constitutes a highly select group of young deaf people. Whatever may be said about differences in physique and environment between deaf and hearing people it would seem that the selection process of Gallaudet College excludes such deaf people as are additionally affected by organic and emotional handicaps. Moreover, among the students would be the ones with the highest intellectual potential of all deaf people. Intelligence tests, standardized on a normal population and usually highly loaded on verbal factors, are recognized as an unsatisfactory means of testing deaf people who are almost universally deficient in verbal language. It seemed, therefore, that by selecting a sample from the college for the deaf and comparing it to a sample of a recognized hearing college, a reasonable solution was found to the problem of equating the groups on intellectual potential. While limited language training in deaf children is self-evident, the lowered performance of deaf college students on verbal achievement tests is noted by writers like Myklebust (6) and others.

2. *Material*

The classification task consisted of 48 white 8 × 11 inch cards on which were glued two geometric figures cut from colored construction paper. From six figures (circle, elipse, crescent, square, rectangle, triangle) and six colors (black, blue, green, red, orange, yellow), two colored figures were systematically selected to exemplify twelve instances each of the following four separate classes: I Form same, Color same (Fs, Cs); II Form same, Color different (Fs, Cd); III Form different, Color same (Fd, Cs); IV Form different, Color different (Fd, Cd).

3. *Procedure*

Two tasks were presented as a group procedure to one class at a time. *S*s were given an answer sheet with 48 numbers on both sides, and were instructed to classify a particular card, placed on a stand in front of the room, by putting one of three letters after the appropriate number. In Task I *E* arbitrarily assigned the letter A as response to the disjunctive class Fs, Cs, or Fs, Cd, the letter B to Fd, Cs, and C to Fd, Cd. Table 1 should clarify the experimental design and the reinforced principles of sorting for all groups.

To facilitate success on Task I three sample cards (Fs, Cd; Fd, Cs; Fd, Cd) with the corresponding response letters were placed before the class and kept in sight throughout this problem. Then *E* placed Card 1 on the

TABLE 1
STIMULUS CATEGORIES AND RESPONSES

Stimulus classes	Response Task I All groups	Task II Group 1	Group 2	Group 3	Group 4
Form same, color same	A*	R	R*	R	R
Form same, color diff.	A*	Q	Q	Q*	Q*
Form diff., color same	B	S*	R*	S	Q*
Form diff., color diff.	C	S*	S	Q*	S

* Disjunctive classes.

stand; after ascertaining that all *S*s had written down a response *E* gave the correct response. No corrections, except for putting an *X* next to a wrong response, no scribbling and no superfluous talking was permitted. After five seconds Card 1 was removed and Card 2 put in its place. Subsequently all 48 cards were presented with corrections. For deaf college students *E* employed a combination of spoken words, signing, and gesturing.

Immediately following completion of Task I, *S*s had to turn over their response sheet, the sample cards of Task I were taken away, and *S*s were instructed to sort on Task II by placing one of three letters Q, R, S after a number. To facilitate understanding of the second task in which no sample cards were used, *E* told the class the correct letter choice for Card 1. While Task I was identical for all groups, Task II differed according to four transfer groups. Apart from the sample cards, response to Card 1, and changed principles of sorting, the procedures for Tasks II and I were identical.

C. RESULTS

On a criterion of nine correct choices in ten successive trials only 10 from a total of 131 hearing and eight of 99 deaf students failed on Task I. Table 2 summarizes the data for all eight groups on Task I and II giving mean error and standard deviation. As for total errors on Task I an analysis yielded an insignificant *F* of 1.72 ($df = 7$) for the between-group variance; but a significant *F* of 6.17 ($df = 2$) demonstrated that the hearing in general

TABLE 2
MEAN ERROR AND STANDARD DEVIATION ON BOTH TASKS

Group	Task I Hearing	Deaf	Task II Hearing	Deaf
1	10.5(6.9)	12.9(5.5)	3.2(3.2)	11.7(7.7)
2	8.4(6.0)	11.1(6.4)	2.7(4.1)	9.3(5.4)
3	8.9(6.2)	9.7(5.3)	5.6(4.7)	9.9(6.9)
4	9.8(4.9)	12.3(8.1)	9.8(4.8)	10.9(6.2)

made less errors than the deaf groups. A second analysis for errors on Task II revealed significant *F*s for between-group (13.56), deafness-hearing (20.00), transfer groups (9.5), leaving only an insignificant *F* of 1.62 for the interaction term. Individual *t*-tests between hearing and deaf transfer groups were significant with the exception of Group 4. Within the hearing transfer groups, Group 3 was found to differ significantly from Group 4 ($t = 3.02$) as well as from Group 2 ($t = 2.40$) and Group 1 ($t = 1.96$).

D. DISCUSSION

Two main results stand out and have a direct bearing on the hypothesis: deaf students in general were poorer than hearing on the transfer tasks and among the transfer tasks there were considerable variations with type of shift. To establish a meaningful explanation of these findings it is essential to have some confidence that Task I did what it was supposed to do, i.e., familiarize *S*s with the components of the classification task. This goal was attained insofar as 95 per cent of deaf and hearing *S*s reached an objective criterion of success. The small overall difference in error scores favoring the hearing does not invalidate the equation of the two groups with respect to reaching criterion, especially since no comparison of any two transfer groups yielded significant differences.

Two classes (Fd, Cs and Fd, Cd) were attained as separate classes with separate responses B and C; the two other classes were attained as a disjunctive class (Fs, Cs, or Fs, Cd). How well an *S* mastered these four classes after success on Task I is a matter for speculation. One would, however, expect that a more developed level of intelligence would be able to separate the components of the disjunctive class and grasp the relationship of the four separate classes better than a less developed intelligence. A perfect mastering of the logical operation of separating and uniting would mean that level of class attainment is the same whether the class is obtained as a separate entity or as component of a combined class. If it were found that intelligent adults approach this ideal, then a computer model for logical human operations would be satisfactory.

In this study level of class attainment was tested by observing deaf and hearing college students on transfer Tasks II. Both groups performed similarly on Task I and differences in their behavior on Task II would manifest differences in level of class attainment. In fact, the deaf groups were poorer on Task II in three of the four tasks. Such a result constitutes evidence for a facilitating influence of language training in classification shift. The help of language, however, should not be understood as giving hearing students better or more specific words for the task at hand. After all, both groups were college students and the words "Color, Form, Different and Same" are quite simple words.

It may be objected that deaf people are supposed to be cognitively rigid in shifting behavior and some evidence for this can be found in the literature (5, 7). The pertinent reports, however, deal consistently with deaf children, often also in an unstructured situation. In our case, all students were told that after Task I there follows a second task possibly different from the first. Finally, the equality of performance between deaf and hearing Transfer Groups 4 would argue against a general rigidity on the part of deaf *S*s.

It seems reasonable to associate more extensive verbal training through ordinary language use with more extensive training in classifying behavior. Thus it could be thought that hearing, more readily than deaf students, would grasp a disjunctive class for what it is and utilize the class knowledge to a greater advantage in a subsequent shift. Such reasoning would be in line with a majority of writers on the intelligence of deaf people (6). Yet in spite of hearing *S*s' supposedly greater level of attainment, manifested particularly in Groups 1 and 2, they were still subject to set and in Group 4 their performance was equal to that of the deaf. The variations within their transfer groups weaken the case for a computer type model. A computer, once having recognized four separate classes combined in any form, would as readily recognize one new combination of these classes as another. Not so with human beings.

A look at Table 1 indicates that a transfer of element theory would predict Group 1 as being the most difficult shift. All three sortings of this group are new in contrast to the other transfer groups who retain at least one sorting class intact from the first task. Yet Group 1 was among the easier shifts. Apparently then even college students did not readily grasp the disjunctive concept as an "either . . . or" combination of two separate classes. How else did they master it?

Bruner (1) points out how in everyday life a constant attempt is made to substitute conjunctive for disjunctive definitions of classes. The following list would seem to express a perfectly logical principle of sorting which does not, however, explicitly grasp the nature of the disjunctive concept.

Task I (All groups) "If *form same, disregard* color."
Task II (Group 1) "If *form* different, *disregard* color."
(Group 2) "If color *same, disregard* form."

(Group 3) "If color different, *disregard* form."
(Group 4) "The combined class differs on one dimension only."

It should be apparent that with such an implicit or explicit grasp of principles, transfer in Group 4 should be most difficult, as in fact it was. Group 3 should come next, as it retained the general verbal outline of principle but changed two terms. Groups 1 and 2 changed only one term and should be easiest. The obtained result followed this reasoning from an implicit verbal principle.

One may also compare transfer of Groups 1, 2 and 3 to reversal shift, transfer of Group 4 to nonreversal shift, and see in their relative difficulty a confirmation of Kendler and Kendler's (4) findings with regard to the ease of reversal over nonreversal shift in college students. The Kendlers' idea of a vertical process seems to be related to what is here termed level of concept attainment.

In the light of the above speculation the former statement about a lower level of concept attainment of deaf students should be given an important modification. Neither deaf nor hearing college students reached the level theoretically obtainable on Task I which would make transfer to any new combination equally easy. Actually among the hearing *S*s only Group 4 had no special benefit from the assumed principle which helped the other transfer groups. For this group knowledge of classes attained through Task I was perhaps the sole source of facilitation. This was the transfer group, it will be remembered, where no difference between the deaf and the hearing students emerged. The deficit of the deaf students was apparently related to less efficient formation of sorting principles which engendered a facilitating set in hearing students.

The question naturally arises whether Group 4 did in fact benefit from the previous Task I, otherwise there would be no point in speaking of transfer. To test for this matter, an additional experiment was performed. A group of 28 students was exposed to Task II of Group 4 without any previous training. Of these only 12 reached a criterion of nine correct choices within 10 trials; the mean error was 20.5 with an *SD* of 6.93. The highly significant difference of 10.7 between the experimental and the control group attests to the positive transfer effect of Task I.

It seems that a whole series of control studies with hearing and deaf people would be in order to clarify the interaction of set and level of attainment. In yet another pilot study *E* observed five students who succeeded first on a training task which required them to sort out the four classes by four separate responses. On a subsequent transfer task similar to Group 4 they made an average of only 3.4 errors ($SD = 1.2$). The difference of 6.8 between this last control group and the experimental group shows clearly the combined effect of set and level on class attainment.

In view of these findings and speculations it would appear that lack of verbal training resulted not so much in a lowered level of concept attainment as in a poorer utilization of set. Insofar as transfer performance of the hearing students was evidently influenced by principles and as verbalization may directly or indirectly influence the formation of principles, deaf students were seen to be less apt in forming and utilizing beneficially such principles. On the one task where verbal set was of no particular help but level of concept attainment seemed the major determining training factor, deaf performed alike to hearing students. This outcome is in agreement with an earlier study (2) where deaf were found to be inferior to hearing children only in the one task where specific language training gave an advantage to hearing children.

The writer is aware that *post hoc* reasoning of this kind can be criticized. Nevertheless, on an *a priori* basis of the four transfer conditions in this experiment, the fourth is least likely to be facilitated by language, and the results both with hearing and deaf students were congruent with this assumption. Convergent findings of two different studies do not constitute a definite confirmation of a major theoretical position. However, in conjunction with other ongoing research, these findings support the thesis that basic cognitive processes, as exemplified in the formation of classes, develop in relative independence from the presence or absence of verbal language.

E. SUMMARY

Four groups of hearing and deaf college students attained four classes in the form of two separate concepts and one disjunctive concept which combined two separate classes. In subsequent transfer tasks in which classes were differently combined the hearing apparently utilized implicit verbal principles to a greater advantage than deaf students on three tasks. The fourth group with no benefit from the principle as a major training factor had a level of concept attainment as in Task I. This group was equal in performance to the deaf group.

Deficit in verbal training apparently influenced formation and utilization of principles of sorting but not level of concept attainment as such. The basic logical operation of class formation was thus seen as relatively independent from verbal language.

REFERENCES

1. BRUNER, J., GOODNOW, J. J., & AUSTIN, G. A. A Study of Thinking. New York: Wiley, 1956.
2. FURTH, H. G. The influence of language on the development of concept formation in deaf children. *J. Abn. & Soc. Psychol.*, 1961, **63**, 386-389.
3. HUNT, D. E., & HOVLAND, C. I. Order of consideration of different types of concepts. *J. Exper. Psychol.*, 1960, **59**, 220-225.
4. KENDLER, H. H., & KENDLER, T. S. Vertical and horizontal processes in problem solving. *Psychol. Rev.*, 1962, **69**, 1-16.
5. MACANDREW, H. Rigidity and isolation: A study of the deaf and the blind. *J. Abn. & Soc. Psychol.*, 1948, **43**, 476-494.
6. MYKLEBUST, H. R. The Psychology of Deafness. New York: Grune & Stratton, 1960.
7. OLERON, P. Conceptual thinking of the deaf. *Amer. Ann. Deaf.*, 1953, **98**, 304-310.
8. PIAGET, J. The Psychology of Intelligence. London: Routledge & Kegan Paul, 1957.

Department of Psychology and Psychiatry
The Catholic University of America
Washington 17, D.C.